UNIVERSITY OF CALIFORNIA
Santa Barbara

Atomic-scale investigations of current and future devices: from nitride-based transistors to quantum computing

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Materials

by

Luke Gordon

Committee in Charge:

Professor Chris G. Van de Walle, Chair

Professor James S. Speck

Professor Chris J. Palmström

Professor David D. Awschalom

September 2014

The Dissertation of
Luke Gordon is approved:

Professor James S. Speck

Professor Chris J. Palmström

Professor David D. Awschalom

Professor Chris G. Van de Walle, Committee Chairperson

September 2014

Atomic-scale investigations of current and future devices: from nitride-based transistors to quantum computing

by

Luke Gordon

To the imminent future

Acknowledgements

I would like to express my recognition of all the people who have made this work possible. First and foremost, I'd like to express my deepest gratitude and thanks to my advisor, Prof. Chris Van de Walle. If I can be described as a "scientist", that is entirely due to his tutelage, advice and guidance over the past 4 years. He cultivates in his students a goals-oriented approach towards research; he instills a fundamental scepticism towards scientific results and always asks "why"; to me, he epitomizes the ideal of falsifiable science, and encourages that ideal through his group meetings, in which the students and post-docs in the group do their best to tear apart any new results, always leaving the presenting student stronger for it. He helped me be confident in my beliefs, and yet willing to admit when I'm wrong. Whatever I am today, it's better than whatever I was before.

Dr. Anderson Janotti, project scientist and all-around awesome person, deserves a special mention. At this point in time, he is the only member of the group, apart from Chris, to have been here since my summer internship in 2009, and is the person whom I've known the longest in Santa Barbara. Over the years, he's been a constant source of advice and encouragement, providing practical advice about running calculations, discussions about ill-formed ideas, and regular pronouncements of "don't worry" (when I'm worrying, or when I'm not). He's also the undisputed master of the grill, cooking

delicious meat, and more meat, and sometimes fish, and always bananas, at our regular beach barbecues.

Next, Democritus and his ancient Greek contemporaries, without whom the concept of an atom would never have been developed, and Albert Einstein, who showed conclusively that Brownian motion could be explained by the random motion of atoms, thereby demonstrating that an atom was a real, measurable thing. Next, I would like to thank Erwin Schrödinger, Werner Heisenberg, and their peers of the quantum mechanics revolution; these brilliant minds developed the theory and framework of quantum mechanics, which underpins our current understanding of atoms. Walter Kohn, who first proposed density functional theory, and all his successors who developed it into the powerful tool that it is today. I'd also like to thank William Shockley, who built the first transistor, Alan Turing, who revolutionized our understanding of computer science, and Richard Feynman, who, in addition to his groundbreaking work in quantum electrodynamics, was also the first proponent of quantum computing. I'm honored to be standing on the shoulders of these giants, and to perhaps contribute in some minuscule way to the understanding they all sought.

In addition, I'd like to thank the many postdocs, past and present, in the Van de Walle group. They have been a source of expertise and advice over the past years, and although many of them have moved on by now, I'd like to express my particular thanks to Dr. Maosheng Miao (my mentor during my summer internship, and an incredibly

patient teacher to a young and ignorant physics student), Dr. Hazem Abu Farsakh (whose project I took over and carried to completion, thanks in no small part to his guidance, in person and abroad), Dr. Audrius Alkauskas (with whom I worked on luminescence of defects in SiC for quantum computing, and who is one of the smartest people I know), and especially Dr. Daniel Steiauf (who has given me patient advice on practically everything I've done, been a great friend, and is overall just a wonderful person). I'd also like to thank Drs. Patrick Rinke, Manos Kioupakis, Poul Moses, Christian Carbogno, Lars Ismer, Khang Hoang, Hartwin Peelaers, Burak Himmetoglu, Minseok Choi and Anindya Roy.

The students in the group have at various points been mentors, mentees, collaborators, sounding boards, and friends. Dr. Jack Lyons was my first mentor in the group, and taught me how to use VASP, and helped me get my first project off the ground. Dr. Joel Varley gave me advice and guidance when starting to study defects as qubits in SiC, a project which he and Dr. Justin Weber, another mentor, initiated. Dr. Qimin Yan gave me general advice, on VASP and on life. Dr. Cyrus Dreyer was my office mate for all but a few months over the past 4 years, and a good friend for all of those years, and patiently enduring my vacuous banter. Finally, Lars Bjaalie, Karthik Krishnaswamy, Wennie Wang, Hiral Taylor, Jimmy Shen, and Patrick McBride have all been good friends, and are all developing into great scientists.

I would also like to thank Profs. Jim Speck, Chris Palmström, and David Awschalom for serving on my committee, and for their help and feedback during the dissertation process, as well as the Awschalom group at large for their collaboration and many fruitful discussions.

Finally, I'd like to thank my friends at UCSB, and my family, who are currently everywhere but UCSB. My parents, Robert and Stephanie Gordon, and my sisters, Maeve and Anna, have always been there to keep me sane, to keep me grounded, and to help me maintain a connection to Ireland.

Thank you all.

Curriculum Vitæ

Luke Gordon

Education

2014	Ph.D. in Materials, University of California, Santa Barbara.
2010	B.A. in Theoretical Physics, Trinity College Dublin, Ireland.

Publications

1. Sukit Limpijumnong, Luke Gordon, Maosheng Miao, Anderson Janotti, and Chris G. Van de Walle.
"Alternative sources of p-type conduction in acceptor doped ZnO", *Applied Physics Letters* **97**, 072112 (2010).

2. Luke Gordon, Maosheng Miao, Srabanti Chowdhury, Masataka Higashiwaki, Umesh K. Mishra, and Chris G. Van de Walle.
"Distributed surface donor states and the two-dimensional electron gas at AlgaN/GaN heterojunctions", *Journal of Physics D: Applied Physics* **43**, 505501 (2010).

3. Anderson Janotti, Lars Bjaalie, Luke Gordon, and Chris G. Van de Walle.
"Controlling the density of the two-dimensional electron gas at the $SrTiO_3$ /$LaAlO_3$ interface", *Physical Review B* **86**, 241108(R) (2012).

4. Luke Gordon, Justin R. Weber, Joel B. Varley, Anderson Janotti, David D. Awschalom, and Chris G. Van de Walle.
"Quantum Computing with Defects", *MRS Bulletin* **38**, 802 (2013).

5. Luke Gordon, John L. Lyons, Anderson Janotti, Chris G. Van de Walle.
"DX Centers in AlN and GaN", *Physical Review B* **97**, 072112 (2014).

6. Luke Gordon, Hazem Abu Farsakh, Anderson Janotti, and Chris G. Van de Walle.
"Hydrogen bonds in Al_2O_3 as dissipative two-level systems in superconducting qubits", *Submitted.*

7. Luke Gordon, Anderson Janotti, and Chris G. Van de Walle.

"Defects as qubits in 3C- and 4H-SiC", *In preparation.*

8. Luke Gordon, Abram Falk, Audrius Alkauskas, Anderson Janotti, and Chris G. Van de Walle.
"Luminescence of defects in SiC", *In preparation.*

9. Luke Gordon, Joel B. Varley, John L. Lyons, Anderson Janotti, and Chris G. Van de Walle.
"Sulfur doping of AlGaN for improved n-type conductivity" *In preparation*

10. Luke Gordon, John L. Lyons, Anderson Janotti, and Chris G. Van de Walle.
"Onset of DX centers in nitride alloys", *In preparation.*

11. Luke Gordon, Anna Sarwe, Daniel Steiauf, Anderson Janotti, and Chris G. Van de Walle.
"The impact of a variable dielectric function in $SrTiO_3$ /$LaAlO_3$ heterostructures", *In preparation.*

12. Burak Himmetoglu, Anderson Janotti, Karthik Krishnaswamy, Lars Bjaalie, Luke Gordon, and Chris G. Van de Walle.
"First-principles Investigation of High Mobility $BaSnO_3$ for Oxide Interfaces", *In preparation.*

13. John L. Lyons, Luke Gordon, Anderson Janotti, and Chris G. Van de Walle.
"Identifying microscopic mechanisms for hole traps in nitride heterostructures", *In preparation.*

Invited talks

2011 The International Conference on Nitride Semiconductors, Glasgow, UK
"DX centers in III-Nitride alloys"

Contributed talks

2011 American Physical Society March Meeting, Dallas, TX
"Hybrid functional calculations of DX centers in GaN and AlN"

2011 The Minerals, Metals, and Materials Society Electronic Materials Conference, Santa Barbara, CA

	"Hybrid functional calculations of DX centers in GaN and AlN"
2012	American Physical Society March Meeting, Boston, MA "Defects as qubits in 3C and 4H-SiC"
2012	Annual Meeting of the California-Nevada Section of the APS, San Luis Obispo, CA "Defects as qubits in 3C and 4H-SiC"
2012	Solid State Lighting and Energy Center Review, Santa Barbara, CA "DX centers in III-Nitride alloys"
2013	American Physical Society March Meeting, Baltimore, MD "Defects as qubits in 3C and 4H-SiC"
2014	American Physical Society March Meeting, Denver, CO "Hydrogen bonds in Al_2O_3 as dissipative two-level systems in superconducting qubits"

Abstract

Atomic-scale investigations of current and future devices: from nitride-based transistors to quantum computing

Luke Gordon

Our era is defined by its technology, and our future is dependent on its continued evolution. Over the past few decades, we have witnessed the expansion of advanced technology into all walks of life and all industries, driven by the exponential increase in the speed and power of semiconductor-based devices.

However, as the length scale of devices reaches the atomic scale, a deep understanding of atomistic theory and its application is increasingly crucial. In order to illustrate the power of an atomistic approach to understanding devices, we will present results and conclusions from three interlinked projects: n-type doping of III-nitride semiconductors, defects for quantum computing, and macroscopic simulations of devices.

First, we will study effective n-type doping of III-nitride semiconductors and their alloys, and analyze the barriers to effective n-type doping of III-nitrides and their alloys. In particular, we will study the formation of DX centers, and predict alloy composition onsets for various III-nitride alloys. In addition, we will perform a comprehensive study of alternative dopants, and provide potential alternative dopants to improve n-type conductivity in AlN and wide-band-gap nitride alloys.

Next, we will discuss how atomic-scale defects can act as a curse for the development of quantum computers by contributing to decoherence at an atomic scale, specifically investigating the effect of two-level state defects (TLS) systems in alumina as a source of decoherence in superconducting qubits based on Josephson junctions; and also as a blessing, by allowing the identification of wholly new qubits in different materials, specifically showing calculations on defects in SiC for quantum computing applications.

Finally, we will provide examples of recent calculations we have performed for devices using macrosopic device simulations, largely in conjunction with first-principles calculations. Specifically, we will discuss the power of using a multi-scale approach to accurately model oxide and nitride-based heterostructures, and thereby illustrate our ability to predict device performance on scales unreachable using a purely first-principles approach.

Professor Chris G. Van de Walle
Dissertation Committee Chair

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> Knowing how things work reduces the effort.
> This is the fundamental principle of
> technology.
>
> Toby Beta

Technology is the application of scientific knowledge for practical purposes, and a device is something made for that purpose. To give an early example, the scientific knowledge that fire gives light was utilized millennia ago to develop devices based on this knowledge: flaming torches were used to explore caves, and bonfires were used to allow humans to function well past the advent of night. Defining device technologies in this way makes many of the more intelligent animals applied scientists: for instance,

chimpanzees who sharpen sticks to hunt monkeys, or who strip stems of their leaves to fish for insects.

However, among Earth's animals, humans are the only ones to have developed this simple process of making knowledge useful into a repeatable science, through our histories and via the generation-by-generation growth of scientific knowledge. We have learned from our history and, at least in scientific terms, we have seldom repeated it.

This dissertation is not a paean to the virtues of science, and to the brave new world that science is ushering in, although it may seem that way at times. Rather it concerns one of the key parts of modern applied science: semiconductor devices, their application, and how we can better understand and improve them through an atomistic approach. Repeating, improving, repeating, improving: this cycle is one that has served humanity well for years, especially when it comes to the explosive advances made in processing power over the past several decades. But the greatest advances have always come from a better understanding of the laws governing our devices, or our galaxy. Greek astronomers, over a period of centuries, developed empirical laws that allowed them to predict with good accuracy the motion of planets, and eclipses of the moon. But without the deeper understanding that came from Newton's, and later Einstein's laws of gravity, their predictions were subject to the caprice of the universe. If something was wrong, they had to add extra models on top of models, trying to fit eccentric orbits with ever more complex systems.

We know a lot about devices, and their operation, but the majority of our knowledge comes from 19th century physical theories of charge transport, coupled with a very approximate usage of quantum mechanics. The thesis of this dissertation is that in order to accurately understand and model future devices, we need to couple our empirical knowledge of devices with an truly atomistic understanding of their operation.

1.1 Device technologies: The past undead, the present known, the future dreamt

In this first section, we will briefly skim the past, present and future of semiconductor device technologies. First, we will start with the development of the first transistors and their exponential increase in power and speed over the ensuing decades. We will then embark on a short survey of the current state-of-the-art, and modern research problems. Finally, we will attempt to foretell the future of devices, and what we need to to do to get there.

1.1.1 The past undead

> The past is never dead. It's not even past.
>
> William Faulkner

The first transistor (a portmanteau of the term transfer resistor) was invented at the famous Bell Labs by John Bardeen, Walter Brattain, and William Shockley, comprising two gold point contacts applied to a crystal of germanium, and the first working silicon transistor was developed at Bell labs by Morris Tanenbaum in 1954. The transistor is a semiconductor device used to amplify and switch electronic signals and electrical power, and is the key component in practically all modern electronics, typically found embedded in integrated circuits along with diodes, resistors, capacitors and other electrical components to produce complete electronic circuits.

Over the ensuing decades, the size of transistors shrunk exponentially, leading Gordon Moore to coin his now-famous Moore's law, which states that the speed of integrated electronic circuits will double every 18 months. Moore's law subsequently became a self-fulfilling prophecy, with engineers at Intel working to maintain the law.

As processor speed exploded, so did its applications. Our research, our consumption of media, our travel: practically everything we see and use on a day-to-day basis has been fundamentally transformed by technology.

1.1.2 The present known

> Because things are the way they are, things will not stay the way they are.
>
> Bertolt Brecht

So where are we now? Intel spends billions of dollars to develop a new fabrication facility, and produces silicon-based integrated circuits with billions of transistors embedded on each chip. New devices such as smartphones require smaller and more efficient processors. The exponential growth in processing speed has slowed, and private and public entities are increasingly looking to expand beyond silicon transistors, and to find whole new paradigms to exploit. New applications such as light-emitting diodes (LEDs) and solar cells require new materials and new device designs; a new demand for high-power, high-frequency electronics has impelled the search for new materials and devices that can meet these lofty goals; and fundamental size limitations on classical bits have led many researchers to try and actualize quantum bits in a quantum computer, in which the effects of quantum mechanics would exponentially improve device operation, rather than limiting it.

1.1.3 The future dreamt

> The best way to predict the future is to invent it.
>
> Alan Kay

The world is becoming increasingly interconnected; billions are being lifted out of poverty and into the middle class; trade barriers are loosening across the world; and scientific knowledge is becoming a public resource. With these great changes come new

markets, and new consumers: solar lighting and cellphone bank transfers in Africa, microlending in India, driverless cars in California—the world of tomorrow is approaching. Technological progress is one of the main drivers of these changes, but is also subject to the demands of the changing consumer.

Through the Materials Genome Project, and other initiatives, an effort is being made to leverage the ability of computer simulations to study new materials systems and devices, before going to the expense of actually making that material or device. Similarly, using atomistic simulations to predict new qubits in materials, as discussed in Chapter 3, is far more cost effective than experimentally testing a range of plausible candidates in different materials.

Overall, as devices get more and more complex, the cost to experimentally test new devices and materials increases, while the computational cost to study said materials and devices drops. We must continue to leverage our atomic understanding of these materials to keep ahead of the curve in a fast-changing world.

1.2 Atomistic theory: From Democritus to Kohn; comprehension and application

In this section, we will provide a brief history of atomistic theory as it grew from pure philosophy to applied science. We will start with the ancients, and their conception

of atoms as an indivisible building block of all matter, which ultimately led to the 19th century development of the periodic table for classifying known elements and predicting unknown elements. We will then move on to the advent of quantum mechanics in the early 20th century, and its transformative implications for atomistic theory. Finally, we will culminate our history with an explanation of the fundamentals of density functional theory (DFT) and its wide-ranging applications in all fields of modern natural science.

1.2.1 Atomism in history: Abstract philosophy to falsifiable science

> Nothing exists except atoms and empty space; everything else is opinion.
>
> Democritus

Atomism began with the ancient Greeks, who theorized that nature consisted of two fundamental principles: indivisible atoms, and empty void. These philosophical atoms, unlike their modern namesakes, came in an infinite variety of shapes and sizes, and could be used as the building blocks of all elements as they were then understood: fire, air, water and earth, and thereby as the indivisible building blocks that made up all matter.

The particles of chemical matter identified in the early 19th century, hydrogen, helium, lithium and the rest, were long thought to be indivisible, and the name "atom"

was therefore used. Until Einstein's characterization of Brownian motion as the result of atoms bouncing against suspended particles in a random walk motion, the physical existence of atoms, or lack thereof, was under debate. By the beginning of the 20th century, atoms were accepted as physical entities, and classified according to the periodic table.

1.2.2 Quantum mechanics: A new dawn

> No one intuitively understands quantum mechanics because all of our experience involves a world of classical phenomena [where, for example, a baseball thrown from pitcher to catcher seems to take just one path,] the one described by Newton's laws of motion. Yet at a microscopic level, the universe behaves quite differently.
>
> Lawrence Krauss

The advent of quantum mechanics heralded a new wave in the understanding of atoms. Although quantum mechanics was born to explain puzzling observations in the blackbody radiation spectrum, it soon became the fundamental theory by which atomic physics was understood. The foundations of quantum mechanics were developed over a

few decades in the early 20th century by Max Planck, Niels Bohr, Werner Heisenberg, Louis de Broglie, Arthur Compton, Albert Einstein, Erwin Schrödinger, Max Born, John von Neumann, Paul Dirac, Enrico Fermi, Wolfgang Pauli, Max von Laue, Freeman Dyson, David Hilbert, Wilhelm Wien, Satyendra Nath Bose, Arnold Sommerfeld and others.

The fundamental equation of quantum mechanics is the Schrödinger equation:

$$i\hbar\frac{\delta}{\delta t}\Psi(r,t) = H\Psi(r,t), \tag{1.1}$$

in which the Hamiltonian H, the operator corresponding to the total energy of the system, generates the time evolution of the system. The wave function $\Psi(r,t)$ determines the state of the system at a given point in time, and the Schrödinger equation can, in theory, be used to solve for any microscopic or macroscopic property of a system. Crucially, and in contrast to classical mechanics, the precision by which certain pairs of physical properties can be determined is fundamentally limited by Heisenberg's uncertainty principle.

1.2.3 Density functional theory: Scaling, improving and applying

> The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.
>
> Paul Dirac

As Dirac observed, quantum mechanics provides a solution to all of our problems in understanding chemical and physical problems, the only problem is in the application. We need simply to solve the many-body Schrödinger equation (below) for a given system, and we know everything there is to know about said system:

$$\begin{aligned}\Bigg[-\frac{1}{2}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i,I}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|} - \sum_I \frac{1}{2M_I}\nabla_I^2 \\ + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|}\Bigg]\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_i\};t) = i\frac{\partial\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_i\};t)}{\partial t},\end{aligned} \tag{1.2}$$

where lower-case subscripts denote a given electron, upper-case subscripts denote a given nucleus, the $\mathbf{r}$'s are the electron positions, the $\mathbf{R}$'s are the nuclear positions, M_I is the nuclear mass, and Z is the nuclear charge. The first term on the left-hand side of Eq. (1.2) is the kinetic energy operator for the electrons, the second term is the coulomb

interaction between electrons, the third term is the potential energy from the electron-nuclei interaction, the fourth term is the kinetic energy of the nuclei, and the final term represents the interactions between the nuclei. The properties of the system are determined by solving for the wave function Ψ.

The solution to Eq. (1.2) will determine everything we might ever need to know about the system: its atomic structure, its electronic structure, its chemical structure, its superconducting T_c, its ionization energies, its color—if it's a measurable quantity, solving the Schrödinger equation will give us the answer.

Unfortunately, solving the Schrödinger equation exactly has turned out to be anything but trivial, and in fact is impossible for all but the simplest of systems, so we need to give up on the idea of ever exactly solving the Schrödinger equation, and instead use approximations to extend our knowledge of the quantum realm to a useful scale. Towards this end, density functional theory (DFT) was postulated in a series of groundbreaking papers in the 1960s.

DFT is a method proposed to make solving the many-body electron problem computationally tractable. Instead of explicitly treating the electron wave functions of the system (which scales exponentially M^N with the number of electrons N), , we instead use the electron density as the defining variable (with just $3N$ variables), which is related

to the many-body wave function through the following relation:

$$n(\mathbf{r}) = \frac{1}{N}\langle\Psi(\{\mathbf{r}_i\})|\sum_{i=1}^{N}\delta(\mathbf{r}-\mathbf{r}_i)|\Psi(\{\mathbf{r}_i\})\rangle. \tag{1.3}$$

Hohenberg and Kohn[1] proved that all (ground- and excited-state) properties of a system in a given external potential are uniquely defined by the ground-state density of the system. In addition, they proved that for any external potential, there exists an *energy functional* of density $E[n(\mathbf{r})]$, whose global minimum is the ground-state energy of the system, and the density that results in this functional being minimized is the exact ground-state density.[1] These theorems put DFT on a theoretically sound footing, and additionally provided a justification for pursuing this method further, due to the vast improvements in efficiency expected from solving in terms of the ground state density, rather than the wave functions.

Carrying out these calculations in practice, however, still requires some way to cast Eq. (1.3) in terms of densities. Some of these practical details were explored in a subsequent publication by Kohn and Sham,[2] in which the many-body-interacting system of electrons is replaced by an auxiliary system of noninteracting particles in an effective potential chosen to reproduce the ground-state density of the many-body system. This effectively allows one to study a system in a single-particle approximation, wherein each particle only feels the effect of a single, effective potential.

In the Kohn-Sham construction, each single-particle wave function has a corresponding single-particle Schrödinger equation (called a Kohn-Sham equation) with an effective potential chosen such that the sum of solutions gives the exact, many-body ground-state density. However, the so-called exchange-correlation functional, V_{XC}, is still undefined, and in fact, there is no known form for the exact V_{XC}. All different functionals used with DFT, such as the local-density approximation (LDA), the generalized gradient approximation (GGA) and various "hybrid" functionals are differentiated by the way in which they define this exchange correlation functional.

In this work we use the hybrid functional developed by Heyd, Scuseria, and Ernzerhof[3] (HSE). Hybrid functionals are characterized by their inclusion of a mixing parameter α which determines the percentage of the Hartree-Fock "exact" exchange energy. In addition, HSE includes a second parameter, σ that screens the long-range part of the Hartree-Fock exchange:

$$E_{\mathrm{XC}} = \alpha E_{\mathrm{X}}^{\mathrm{HF,SR}}(\sigma) + (1-\alpha)E_{\mathrm{X}}^{\mathrm{GGA,SR}}(\sigma) + E_{\mathrm{X}}^{\mathrm{GGA,LR}}(\sigma) + E_{\mathrm{C}}^{\mathrm{GGA}}, \tag{1.4}$$

where the Hartree-Fock exchange is now replacing only the short-range exchange of the GGA functional.

Whether or not hybrids truly count as "first-principles", given that they include an empirically determined mixing parameter is a question best left to better minds.[4] For this work, we will simply say that in general, hybrids provide an improved description of

the electronic structure of semiconductors and insulators, especially producing far more accurate band gaps and charge transition levels.[5,6] The gap in a given material has an approximately linear dependence on the α parameter in Eq. (1.4), and we use different mixing parameters in this work, chosen to fit the experimentally determined band gap.

1.3 Reconciliation: On theory and practice

> In theory, theory is the same as practice. In practice, it rarely is.
>
> Yogi Berra

In this section we will describe how we can apply first-principles theory to understand and predict experiment from an atomic scale and how we can use that first-principles knowledge to extend the scope of our studies to macroscopic, device-scale properties.

1.3.1 Applying first-principles theory to understand and predict experiment

> Make everything as simple as possible, but no simpler.
>
> Albert Einstein

As described above, we use the HSE functional to perform calculations on systems with ∼100 atoms. But what can we actually calculate with this functional, and why is it useful? This section will provide a whirlwind tour of various different properties we can calculate with HSE, focusing on those for which specific examples are provided in the succeeding chapters.

We can calculate the formation energy—i.e., the likelihood of a defect in a given charge state incorporating in a crystal.[7] For the specific example of an impurity X incorporating in a material M, the formation energy $[\mathrm{E^f}(\mathrm{X}^q]$ in M is given by:

$$E^{\mathrm{f}}(\mathrm{X}^q) = E_{\mathrm{tot}}(\mathrm{X}^q) - E_{\mathrm{tot}}(\mathrm{M}) - \mu_{\mathrm{X}} + qE_{\mathrm{F}} + \Delta^q, \tag{1.5}$$

where $E_{\mathrm{tot}}(\mathrm{M})$ is the total energy of the material M with no impurity, and $E_{\mathrm{tot}}(\mathrm{X}^q)$ is the total energy of M with a defect in charge state q. μ_{X} is the chemical potential of the impurity; in addition, the chemical potential(s) of any atoms removed to make room for the impurity must be added to the formation energy. Solubility-limiting phases (i.e., the chemical potential at which a given impurity is more likely to form another phase, rather than substituting on the lattice site) need also to be taken into account, and this will set an upper bound on the chemical potential of the impurity. E_{F} is the Fermi level, which is referenced to the valence-band maximum (VBM). One issue that arises when studying systems with charged defects is the charged defect "seeing" its reflection in the periodic supercell, and introducing spurious charge state contributions to the total energy. We

address this using a finite-size correction Δ^q for charged defects, calculated using the procedure outlined in Refs. 8 and 9. A more specific example of this calculation is given in Section 2.2.

We can calculate transition energies between different charge states, between defect and band states and between differing defect states, allowing us to predict and interpret absorption and emission energies, as illustrated in Chapter 3. We can calculate the specific form of photoluminescence spectra, including the effect of phonons, to a high degree of accuracy, and we can calculate hyperfine constants for all defects.

In general, first-principles calculations can be used to describe and explain a wide variety of physical phenomena and parameters, and have recently been employed by our group to study such diverse subjects as polarization in nitride materials, band offsets at complex oxide interfaces, and why ZnO light emitters based on *p*-*n* junctions are impossible.

1.3.2 Extending first-principles results to device-scale calculations

> All exact science is dominated by the idea of approximation.
>
> Bertrand Russell

Newton's laws of gravitational motion are an approximation to general relativity. Classical laws are a macroscopic approximation to quantum mechanical laws. Density functional theory is an approximation to the full many-body Schrödinger equation. All known laws are presumably an approximation to some grand unified theory which describes everything (unless they're not, and there isn't). In this section, and in Chapter 4 of this thesis, we describe how we can solve the coupled Schrödinger-Poisson equations, using first-principles and experimental input, to study the operation of devices at the macroscopic scale. All calculations shown later were performed using the nextnano^3 (Refs. 10,11) package, although in many cases modifications to the source code were made, or additional scripts were written, to enhance its functionality.

The Poisson equation describes the electrostatics within the device, and is defined as follows:

$$\nabla\epsilon(x)\nabla\phi(x) = -\rho(x), \tag{1.6}$$

where ϵ is the dielectric constant, ρ is the charge density, and ϕ is the electrostatic potential.

The charge density ρ within a semiconductor device is given by:

$$\rho(x) = e[-n(x) + p(x) + N_D^+(x) - N_A^-(x) + \rho_{fix}(x)], \tag{1.7}$$

where e is the electron charge, n and p are the electron and hole densities, and N_D^+ and N_A^- are the ionized donor and acceptor concentrations. This equation is first discretized

using a finite differences method, then solved numerically with an iterative Newton-Raphson scheme.

The ionized shallow donor and acceptor densities are given by:

$$N_D^+(x) = \Sigma \frac{N_{D,i}(x)}{1 + g \cdot exp(E_{F,n}(x) - E_{D,i}(x)/k_B T)}, \tag{1.8}$$

and

$$N_A^-(x) = \Sigma \frac{N_{A,i}(x)}{1 + g \cdot exp(E_{A,n}(x) - E_{F,i}(x)/k_B T)}, \tag{1.9}$$

where k_B is the Boltzmann constant, g is the degeneracy factor, T is the temperature, $E_{F,n}$ and $E_{F,p}$ are the electron and hole Fermi levels, $N_{D,i}$ and $N_{A,i}$ are the donor and acceptor concentrations, and $E_{D,i}$ and $E_{A,i}$ are the energies of the neutral donor and acceptor impurities, respectively.

Of course, in addition to solving the purely classical Poisson equation, we must solve the Schrödinger equation self-consistently with respect to the potential energy $V(x)$:

$$(-\frac{\hbar^2}{2} \nabla M(x) \nabla + V(x)) \Psi_n(x) = E_n \Psi_n(x), \tag{1.10}$$

where $M(x)$ is the effective mass tensor and Ψ_x is the wavefunction. In all simulations presented, we solve the one-dimenional Schrödinger equation, given by:

$$(-\frac{\hbar^2}{2} \frac{\delta}{\delta x} (\frac{1}{m(x)} \frac{\delta}{\delta x}) + V(x)) \Psi_n(x) = E_n \Psi_n(x), \tag{1.11}$$

where $m(x)$ is the effective mass tensor component along the x direction. For a specific example of this, please refer to Section 4.3.1.1.

Chapter 4 provides several examples of using Schrödinger-Poisson simulations to study systems such as oxide and nitride heterostructures. By using such calculations, and combining them with our atomistic knowledge of the operation of devices, we can understand and predict device operation at a scale unreachable by purely first-principles calculations.

1.4 The power and the glory of a first-principles approach: recent investigations

This section will provide a brief preview of the subject of the remaining chapters in this thesis. Specifically, we will discuss how we can use first-principles calculations to study n-type doping of III-nitrides and their alloys, and how we can improve the n-type conductivity in wide-band-gap nitrides through alternative dopants. In addition, we will discuss the role of defects in quantum computing: their deleterious effects on superconducting qubits based on Josephson junctions, and their potential applications in defects based on qubits, modeled after the nitrogen-vacancy center (NV) in diamond. Finally, we will show how first-principles results can be extended to the macroscopic scale, by presenting explicit macroscopic simulations of oxide and nitride device heterostructures.

1.4.1 Effective n-type doping of III-nitrides and their alloys

> One group of riders doped, the others alongside them racing clean. You can work out for yourselves which group was fastest.
>
> David Millar

In Chapter 2, we will exhibit first-principles calculations on defects in III-nitride materials and their alloys. The III-nitride semiconductors (AlN, GaN and InN) are among the most widely-researched modern semiconductors: they have already been commercialized in power electronics and optoelectronic devices, and are promising candidates for ultraviolet light-emitting diodes, biochemical-sensing, and spintronics applications.

For wide-band-gap III-nitrides such as high-Al-content AlGaN and AlInN, doping has been challenging. In Chapter 2, we will analyze the barriers to effective n-type doping of III-nitrides and their alloys, particularly the formation of DX centers. We will perform a comprehensive study of alternative dopants, and provide potential solutions to the n-type doping of AlN and wide-band-gap nitride alloys.

1.4.2 Defects in quantum computing: A double-edged sword

> Just because something doesn't do what you planned it to do doesn't mean it's useless.
>
> Thomas Edison

In Chapter 3, we will study the ambivalent role of defects in quantum computing.[12] We will first study how atomic-scale defects can act as a negative factor in the development of quantum computers by contributing to decoherence at an atomic scale; specifically, we will present work on the origin of dissipative two-level systems in alumina, a common dielectric material in superconducting qubits based on Josephson junctions. Next, we will discuss how defects can be leveraged as solid-state qubits, using the specific example of the NV-center in diamond, and we will describe how we can use DFT to identify wholly new qubits in different materials, such as SiC.

1.4.3 Beyond the cubic nanometer: macroscopic simulations of devices

> It doesn't stop being magic just because you know how it works.
>
> Terry Pratchett

In Chapter 4, we will use first-principles and, in some cases, experimental input, to extend our calculations to a larger, macroscopic length scale, and discuss how a multi-scale approach incorporating first-principles input and Schrödinger-Poisson calculations allow us to accurately model oxide and nitride-based heterostructures, We will first present work on the origin and properties of the two-dimensional electron gas (2DEG) at the $SrTiO_3/LaAlO_3$ (STO/LAO) interface. In addition, we will present calculations on AlGaN/GaN high-electron-mobility transistors (HEMTs), and the role of surface states and microscopic hole traps in determining the properties of the 2DEG. We will thereby illustrate our ability to predict device performance on scales unreachable using a purely first-principles approach.

1.5 Methods: The practical stuff

> It's supposed to be automatic, but actually
> you have to push this button.
>
> John Brunner

In the following section, we will provide notes and explanation of a few of the practical aspects of running and interpreting first-principles and device calculations.

1.5.1 Performing and interpreting first-principles calculations

> I know there's a proverb that says "To err is human", but a human error is nothing to what a computer can do if it tries.
>
> Agatha Christie

In the above discussion of DFT and HSE, we described the methods by which we can use first principles calculations to explain and predict experiment. Here we will make some comments on the practical aspects of running these calculations. All DFT calculations were carried out in the Vienna *ab initio* simulations package (VASP).[13] The exchange-correlation (XC) functionals used were PBE-GGA[14] and HSE.[3]

First, in order to solve the Kohn-Sham equation numerically, we must choose a basis set to parametrize the wave functions. In all calculations we show, we have expanded the wave function in plane waves. This allows for simple convergence checks, as the accuracy of the plane-wave expansion increases monotically in concert with the highest-energy plane waves included in the basis set. This is specified by a system-dependent energy cutoff parameter—for instance, we use an energy cutoff of 300 eV for all III-nitride calculations shown, and an energy cutoff of 400 eV for SiC calculations. One issue with using a plane-wave basis set is that the number of plane waves requires to accurately describe a rapidly varying potential, such as that describing the core electrons in an

atom, is prohibitively large. This is addressed through the use of pseudopotentials, which replace the wave functions of the core electrons with a more smoothly varying potential, that still accurately describes the long-range behavior of the full-electron potential. In our calculations we use the Projector Augmented Wave (PAW)[15] method to generate such pseudopotentials.

All of our calculations are performed using periodic boundary conditions in all directions, allowing the simulation of macroscopic solids with a relatively small number of atoms, as long as that solid has well-defined periodicity (i.e., has long-range order). In this text, calculations are presented for supercells with atoms numbering between 64 and 144.

We must choose k-points in the reciprocal space that accurately give us the total energy for the system in question. It was shown, first by Baldereschi[16] and later by Chadi and Cohen[17] and Monkhorst and Pack[18] that for a given number of k-points, we can take advantage of the symmetry of the system to determine which k-points will give us the best approximation to a full integration over the BZ. In the majority of cases shown here, our supercell is sufficiently large that using a regular 2x2x2 grid of k-points is enough to give us the required accuracy in the total energy.

For cases where energy differences need to be calculated to a very high degree of accuracy, a denser grid of k-points may be needed. In particular, when calculating absorption

and emission energies of defect spin states using constrained DFT (see Section 3.3), we must take care that the system is completely converged with respect to k-points.

1.5.2 Schrödinger-Poisson calculations

> Truth is much too complicated to allow anything but approximations.
>
> John Von Neumann

As noted above, all Schrödinger-Poisson calculations were performed using the nextnano[3] software. A few practical notes are below, and full details are provided in Chapter 4.

Choosing the spatial grid point density correctly is essential. Too dense, and you're wasting time, too sparse and you're missing physics. Typically, if the change in carrier concentration over a given distance is minimal, one does not need many grid points. Conversely, if the change in carrier concentration is large (such as at the STO/LAO interface or the AlGaN/GaN interface), far more grid points are required. A good method for checking convergence is to increase the number of grid points until the change in integrated carrier density is minute. It is usually less computationally intensive to define different regions with different grid point densities, depending on the expected change in potential over said region.

In the case of the STO/LAO calculations, the macroscopic simulations were not "aware" of the polar/non-polar interface between STO and LAO (indeed, parameters for all complex oxides need to be entered manually, and approximated as zinc-blende materials). To study the effect of a 2DEG at the interface, a sheet of carriers was introduced at the STO/LAO boundary, with the volume charge density picked to correspond to the expected $3.3\text{x}10^{14}$ cm^{-2} carrier density.

1.6 Conclusions

> Your present circumstances don't determine where you can go; they merely determine where you start.
>
> Nido Qubein

Technological progress through the application of scientific knowledge has been one of the great driving forces behind human advancement for millennia, and it has become particularly apparent over the past several decades, as computers, the internet, and smartphones have expanded into every industry. In order to continue this rapid advancement of technology, we must fully understand its workings from an atomistic perspective, and apply our knowledge to maintain and accelerate the development of new devices and technologies to keep up with the dynamism of the world. Over this first

chapter, we have stressed the importance of an atomistic approach to devices; over the next three chapters, we will provide specific examples of our calculations for devices to directly illustrate the power of an atomistic approach to understanding device operation, both for improving existing device performance, and for developing whole new types of devices.

Chapter 2

Effective n-type doping of III-nitrides and their alloys

> If my mother knew I did this for a living,
>
> she'd kill me. She thinks I'm *selling* dope.
>
> Henny Youngman

2.1 Introduction

In advancing today's technology towards more powerful and efficient solutions, we must look beyond Si and other traditional semiconductors towards new materials and new device constructions. New global markets for wireless communication, power-

efficient lighting, and sustainable energy generation require innovations at all parts of the spectrum, and keeping up with this demand necessitates the developments of new component technologies such as power transmitters and receivers, and new optoelectronic devices that can more efficiently absorb and emit light for energy generation, and light emission purposes. The group-III nitride semiconductors and their alloys[19] are among the most exciting possibilities for next-generation solar cells and lighting applications, with many important commercial applications in optoelectronic devices that operate in the visible spectrum. They are promising candidates for fabricating ultraviolet light-emitting diodes (UV LEDs) and laser diodes, which could replace gas lasers and mercury lamps for applications such as air and water purification, disinfection and high-resolution photolithography.[20]

To achieve high-efficiency UV LEDs and lasers, high Al-content AlGaN alloys grown, preferentially, on high-quality bulk AlN substrates will be required,[21] along with a better understanding of the role of defects and impurities in AlN. All devices require doping, and achieving highly conducting n-type AlN or high-Al-content nitride alloys has been challenging. For instance, oxygen is known to be a common impurity in nitride semiconductors, although its behavior is still under active investigation for AlN and $Al_xGa_{1-x}N$ alloys.[22,23] Oxygen, which is a shallow donor in GaN,[24] ceases to produce n-type conductivity at moderate Al concentrations,[25–27] and any nitride alloy with a significant Al concentration will have similar doping difficulties. Silicon is the most commonly used

n-type dopant in all nitride alloys,[19,28] but for AlN and high Al-content alloys, it is difficult to achieve conductivities on a par with lower-band-gap III-nitrides.[29] This is due to three main factors: an increase in the donor ionization energy,[30] a decrease in formation energy of compensating acceptors such as Al vacancies[31,32] and carbon acceptors,[24] and the occurrence of a DX transition.[27,33,34]

In this chapter, we will analyze the barriers to effective n-type doping of III-nitrides and their alloys, particularly the formation of DX centers (Section 2.3). We will then perform a comprehensive study of alternative dopants, and provide potential solutions to the n-type doping of AlN and wide-band-gap nitride alloys (Section 2.4). All results on oxygen, silicon, and germanium in AlN and GaN shown in this chapter were published in Physical Review B.[27] Manuscripts on all other results are in preparation.[35,36]

2.2 Methodology and materials parameters of the III-nitrides

We fitted the mixing parameters of AlN, GaN and InN, choosing 33%, 31% and 25% respectively, to reproduce the experimental band gaps. The lattice parameters and band gaps of wurtzite AlN, GaN and InN calculated using PBE and HSE are listed in Table 2.1 and compared with experimental values from Ref. 37. Both PBE and HSE give

Table 2.1: Lattice parameters (a and c/a ratio) and band gap (E_g) of wurtzite AlN, GaN and InN calculated using the PBE and HSE functionals. Experimental values from Ref. 37 are listed for comparison.

Material	Property	PBE	HSE	Exp.
	a (Å)	3.13	3.10	3.11
AlN	c/a	1.60	1.60	1.60
	E_{g} (eV)	4.09	6.10	6.19
	a (Å)	3.22	3.20	3.19
GaN	c/a	1.63	1.62	1.63
	E_{g} (eV)	2.06	3.50	3.50
	a (Å)	3.58	3.58	3.54
InN	c/a	1.62	1.61	1.61
	E_{g} (eV)	0.00	0.65	0.70

consistently accurate structural parameters for AlN, GaN and InN, but PBE severely underestimates the band gaps.

The likelihood of a defect incorporating in a crystal is given by its formation energy.[7] For the case of a sulfur impurity on the nitrogen substitutional site, the formation energy $[\mathrm{E^f}(\mathrm{S_N^q})]$ is given by:

$$E^{\mathrm{f}}(\mathrm{S_N^q}) = E_{\mathrm{tot}}(\mathrm{S_N^q}) - E_{\mathrm{tot}}(\mathrm{AlN}) - \mu_{\mathrm{S}} + \mu_{\mathrm{N}} + qE_{\mathrm{F}} + \Delta^q, \tag{2.1}$$

where $E_{\rm tot}$(AlN) is the total energy of the pristine AlN supercell and $E_{\rm tot}(S_N^q)$ is the total energy of the structure containing the defect in charge state q. $\mu_{\rm N}$ and $\mu_{\rm S}$ are the chemical potentials of N and S. The chemical potentials μ_{Al} and $\mu_{\rm N}$ can vary over ranges limited by the enthalpy of formation of AlN, which is calculated to be -3.36eV. Therefore, we need to consider both N-rich and Al-rich conditions. Solubility-limiting phases need also to be taken into account, and this will set an upper bound on the chemical potential of the impurity. In the case of AlN with P, we use AlP as a solubility limiting phase, for As we use AlAs, for Sb we use AlSb, for F we use AlF, for S we use Al_2S_3, and for Se we use Al_2Se_3. No stable solubility limiting phases were found for Sn or Pb. Δ^q is the finite-size correction for charged defects calculated using the procedure outlined in Refs. 8 and 9, and $E_{\rm F}$ is the Fermi level, which is referenced to the valence-band maximum (VBM). Analogous equations apply to other impurities.

2.3 DX centers in III-nitrides and their alloys

The formation of DX centers has been proposed to explain experimental results obtained from electron paramagnetic resonance (EPR) studies and conductivity measurements in AlN and AlGaN alloys.[38] DX centers occur when an impurity expected to act as a shallow donor instead undergoes a large, bond-rupturing displacement, and traps electrons to become a deep acceptor. The prototype DX center is Si in GaAs and

AlGaAs.[39,40] In AlGaN, conductivity measurements as a function of composition found sharply decreasing conductivity with increasing Al content, along with the presence of oxygen impurities in the samples.[38] This behavior was explained by proposing that oxygen was the dominant impurity in these samples, and self-compensated by forming DX centers for Al content greater than 0.27. However, other researchers have observed n-type conductivity in $Al_xGa_{1-x}N$ alloys for x greater than 0.6 without intentional doping,[41] which they ascribed to the presence of unintentional oxygen donors that acted as the source of n-type conductivity at these alloy concentrations. Pophristic *et al.*[42] also observed an increase in conductivity when the oxygen concentration was increased in Si-doped $Al_{0.6}Ga_{0.4}N$ alloys.

Silicon is commonly used as an intentional n-type dopant[43] in AlN and AlGaN alloys. The Si donor has been predicted to form DX centers in $Al_xGa_{1-x}N$, but predictions for the onset of DX behavior as a function of Al concentration vary widely.[25,44,45] Electron paramagnetic resonance (EPR) measurements also indicate that Si forms a DX center in AlN.[46,47] Recent results predict the formation of DX centers with a (+/−) transition level located ∼0.14 eV below the conduction-band minimum (CBM) in AlN.[46] The authors speculate that this state is close enough to the CBM for Si to behave as a shallow donor in AlN. Very recent experiments from the same group seem to confirm this, indicating from temperature-dependent EPR measurements that Si is a DX center in $Al_{0.77}Ga_{0.23}N$, but behaves as a shallow donor due to a small separation of just 3

meV between the neutral E_d and the lower-lying negative state E_{DX}.[48] On the other hand, conductivity measurements suggest that Si impurities form DX centers for Al concentrations as low as 0.5 in AlGaN.[49]

The seeming lack of consistency between the various results that have been reported in the literature indicates that there is still a need for fundamental studies that address the atomic and electronic structure of the oxygen and silicon impurities, as well as an exploration of other potential donors.

The formation of a DX center leads to the self-compensation of a shallow donor through the formation of an acceptor state. This reaction can be written as:[44]

$$2d^0 \rightarrow d^+ + \mathrm{DX}^-, \tag{2.2}$$

where d^0 is the donor impurity in the neutral charge state, d^+ is the positively charged donor, and DX^- represents the impurity in the negatively charged DX configuration. The effective correlation energy U associated with this reaction can be defined by:

$$U = E^+ + E^- - 2E^0, \tag{2.3}$$

where E^+ and E^0 are the formation energies of the donor impurity in the positive and neutral charge states, and E^- is the formation energy of the negatively charged DX configuration. It has been traditionally assumed that if U is negative, this indicates a stable DX center. A simpler condition for DX-center stability is that the (+/–)

transition level lies below the conduction band, and this is used as the criterion for all systems studied below.

Using density functional theory (DFT) based on the local-density approximation (LDA) or the generalized gradient approximation (GGA), previous theoretical studies have examined the properties of O, Si and other common impurities in nitride semiconductors.[45] It has been found that both O and Si act as shallow donors in GaN,[25,44] but the results for DX behavior in AlGaN alloys have varied widely. Park and Chadi[44] predicted the onset of DX behavior for Si to occur at x >0.24 in $Al_xGa_{1-x}N$, while Boguslawski and Bernholc,[45] using LDA, predicted that Si-related DX centers are stable at $x > 0.60$. Yet another study predicted that Si_{Al} remains a shallow effective-mass donor in AlGaN over the whole composition range.[25] For O_N in AlGaN, the results have been more consistent. Based on GGA calculations, Van de Walle predicted the formation of O-related DX centers to occur for $x \geq 0.46$,[25] while Park and Chadi predicted DX centers to form for $x \geq 0.20$.[44]

In this section, we first present results on the formation and stability of various DX centers in AlN and GaN; next, we provide predictions as to the likely onset of DX behavior in the nitride alloys AlGaN and AlInN.

2.3.1 Stability of DX centers in AlN and GaN

2.3.1.1 Oxygen in AlN and GaN

Formation energies for the positive, neutral, and negative charge states of O_N are presented in Fig. 2.1 for both AlN and GaN. The results are presented for N-rich conditions; the choice of μ_N affects the absolute values of the formation energies, but not the position of transition levels in the band gap. As Fig. 2.1(b) shows, O_N is a shallow donor in GaN, as the positive charge state is most stable for all Fermi levels within the band gap. We therefore set the (+/0) transition level equal to the CBM. A DX configuration can be found that is locally stable in GaN, however, the (+/–) transition occurs at 0.59 eV above the CBM. The atomic configurations of O_N and the DX state in GaN are very similar in nature to those for AlN, shown in Fig. 2.2. In AlN, O_N is stable in a DX configuration when the Fermi level is above the (+/–) transition level, located at 0.64 eV below the CBM. As seen in Fig. 2.1(a), O_N exhibits negative-U behavior, with $U = -0.36$ eV. The lowest-energy configuration of O_N^+ is shown in Fig. 2.2(a). In the positive charge state, the O atom sits on the substitutional N site, with breathing-mode relaxations of the nearest-neighbor Al atoms corresponding to a 4% elongation of the O-Al bond length with respect to the Al-N bulk bond length of 1.90 Å. The lowest-energy form of the DX configuration for O_N^- is shown in Fig. 2.2(b). The O atom is displaced along the basal Al-O bond axis (as opposed to the Al-O bond along the c-axis).

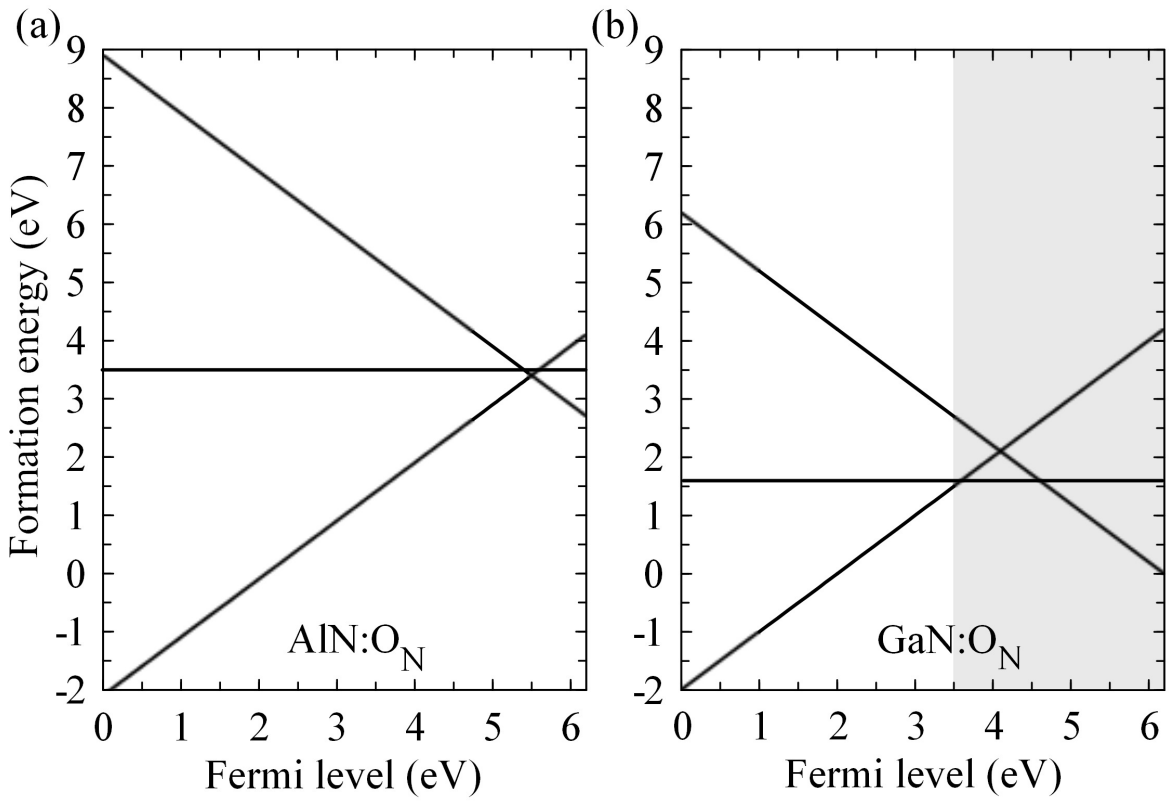

Figure 2.1: Formation energy as a function of Fermi level for O_N in (a) AlN and (b) GaN, for N-rich conditions. For GaN, Fermi levels above the CBM (at 3.5 eV) are indicated by the shaded area.

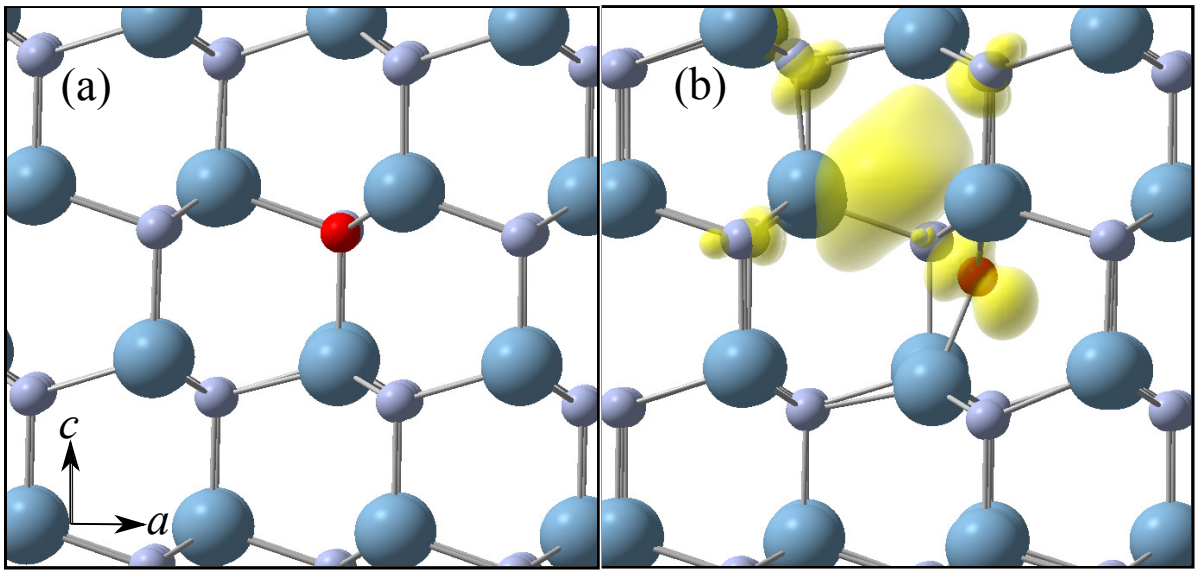

Figure 2.2: Atomic structures of (a) O_N^+ and (b) O_N^- in the most stable DX-center configuration in AlN. The oxygen atom (small red sphere) sits on the substitutional site normally occupied by nitrogen (small light blue spheres); aluminum atoms are denoted by large blue spheres. The yellow lobes in (b) indicate the charge density associated with the occupied DX state in the band gap, with the isosurfaces set to 5% of the maximum.

The bond lengths to its three Al neighbors are 1.82 Å, 1.82 Å and 1.75 Å, contracting compared with the Al-N bulk bond length.

We tested a number of different possible DX configurations for O_N^-, including the configurations explored in previous work.[44] Most previous calculations found that for the stable oxygen DX configuration in AlN, the axial bond was broken, and the oxygen was displaced along the c-axis.[25,44] The lowest energy configuration for O_N^- in the present work has an energy approximately 0.2 eV lower than the axial DX state. The local lattice relaxation of the lowest-energy configuration and the charge density of the acceptor state in the gap are shown in Fig. 2.2(b).

Based on the values for the (+/–) level for O in GaN and AlN, we predict the DX formation to occur for x=0.61 in $Al_xGa_{1-x}N$ (see Table 2.2), which is the Al concentration where the (+/–) transition will be resonant with the CBM in the alloy. For Al concentrations higher than this value, O_N^- in AlGaN will be stable in the band gap for Fermi levels above the (+/–) level, and self-compensation of the oxygen donor will occur. If the (+/–) level is sufficiently far below the CBM, the Fermi level will essentially be pinned at the (+/–) level, and the concentration of free electrons will be much smaller than the impurity concentration. Assuming an effective mass of 0.3 m_0 for AlGaN, a carrier concentration of 10^{15} cm^{-3} corresponds to a Fermi level of approximately 0.20 eV below the CBM. Our numbers in Table 2.2 imply that up to an Al concentration of 75% in the alloy, oxygen might still give rise to an observable level of n-type doping, since

Table 2.2: Properties of donor impurities in InN, GaN, and AlN. (+/–) indicates the transition level between the positive and negative charge state of the impurity, expressed relative to the CBM (in eV). *DX* onset is the Al concentration where the (+/–) transition level coincides with the CBM. Band bowing parameters of 0.7 eV and 2.5 eV[50] are used for AlGaN and AlInN, respectively.

Impurity	AlN (+/–)	GaN (+/–)	InN(+/–)	*DX* onset (AlGaN)	*DX* onset (AlInN)
O	-0.64	0.59	1.95	0.61	0.86
Si	-0.15	1.65	3.33	0.94	0.97

the (+/–) level would be less than 0.2 eV below the alloy CBM. This may explain why increases in conductivity due to oxygen may still be found in $Al_xGa_{1-x}N$ alloys with x as large as 0.6.[41,42] Also, since the presence of other dopants (or compensating centers) affects the relative position of the Fermi level with respect to the oxygen (+/–) level, the onset of *DX* behavior may not be clear cut in experimental measurements.

2.3.1.2 Silicon in AlN and GaN

The formation energies of Si_{Al} are shown in Fig. 2.3(a). We find that Si_{Al} forms a *DX* center in AlN, with U=–0.30 eV. The (+/–) transition level occurs at 0.15 eV below the CBM, much closer than for O_N.

The atomic configurations for the positive and negative charge states of Si_{Al} are shown in Fig. 2.4. For Si^+_{Al}, the silicon occupies the cation site, with relaxations of the nearest neighbor nitrogen atoms corresponding to a 6% contraction of the Si-N bonds with respect to the bulk Al-N bond length. The lowest-energy *DX* configuration for

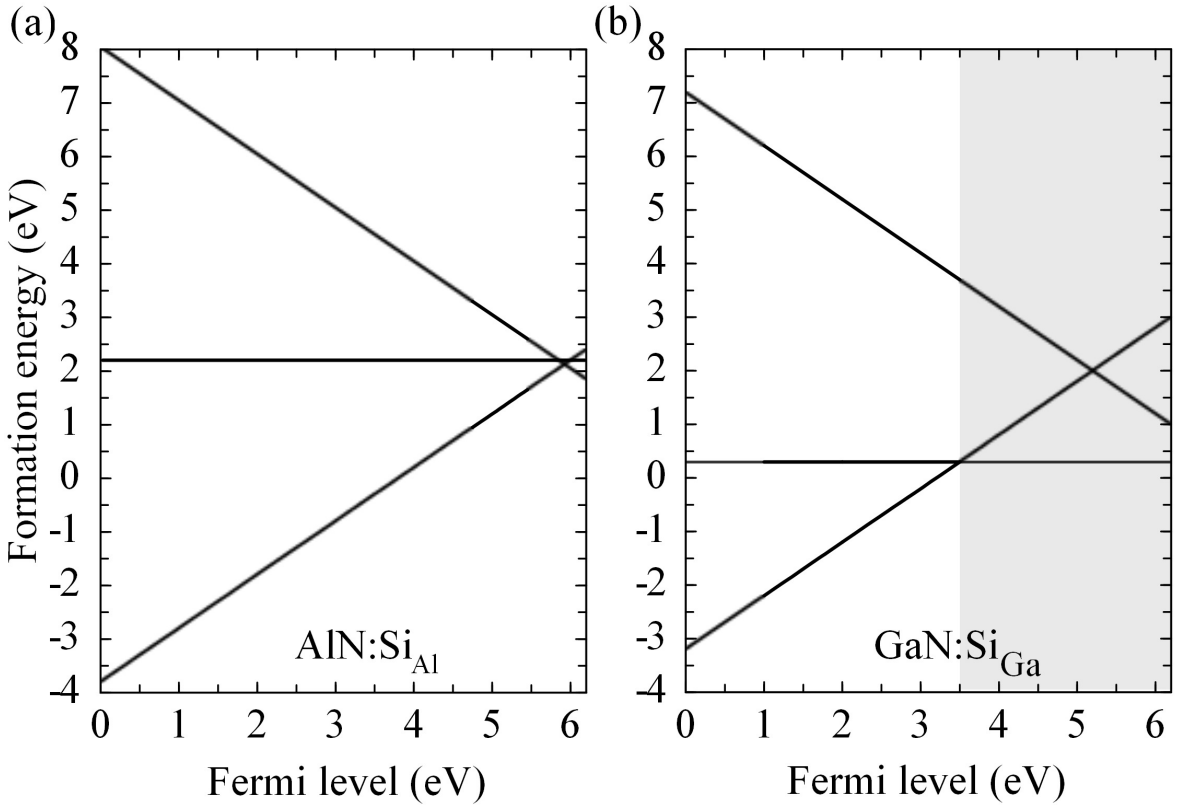

Figure 2.3: Formation energy versus Fermi level for Si on the cation substitutional site in (a) AlN and (b) GaN, for N-rich conditions. In GaN, Fermi levels above the CBM are indicated by the shaded area.

Si_{Al}^{-} is shown in Fig. 2.4(b). The Si atom stays close to the substitutional site, with a contraction of 2% of the surrounding N atoms; the axial bond along the *c*-axis is broken due to a large displacement of the N atom. The charge density of the *DX*-induced occupied state in the gap is also shown in Fig. 2.4(b), indicating its localized, broken-bond character. Our results differ from the "α-broken-bond" configuration of Park *et al.* (Ref. 44), in which bond breaking along the *c*-axis occurs due to a large displacement of the Si (as opposed to the N) atom. Such a configuration is not stable in our calculations. Recent calculations on O and Si *DX* centers in AlN using HSE[51] found similar lowest-energy configurations to those found in the present work. However, the exact positions of the (+/–) levels differ by 0.1-0.2 eV from the values calculated here, likely due to the different mixing parameter and different finite-size corrections.

The calculated formation energies for Si_{Ga} in GaN are shown in Fig. 2.3(b). Silicon is a shallow donor in GaN, with Si_{Ga}^{+} stable for all Fermi levels inside the band gap. We are able to stabilize a *DX* configuration for Si_{Ga} in GaN, but the (+/–) transition occurs 1.65 eV above the CBM. For Si_{Ga}^{+}, the silicon occupies the cation site, with relaxations of the nearest neighbor nitrogen atoms corresponding to a 9% contraction of the Si-N bonds with respect to the bulk Ga-N bond length.

To predict the onset Al concentration for the formation of the Si_{cation} *DX* center in AlGaN alloys, we again interpolate between the (+/–) transition levels of pure AlN and GaN, including the effects of band bowing; we thus predict the onset to occur for x=0.94

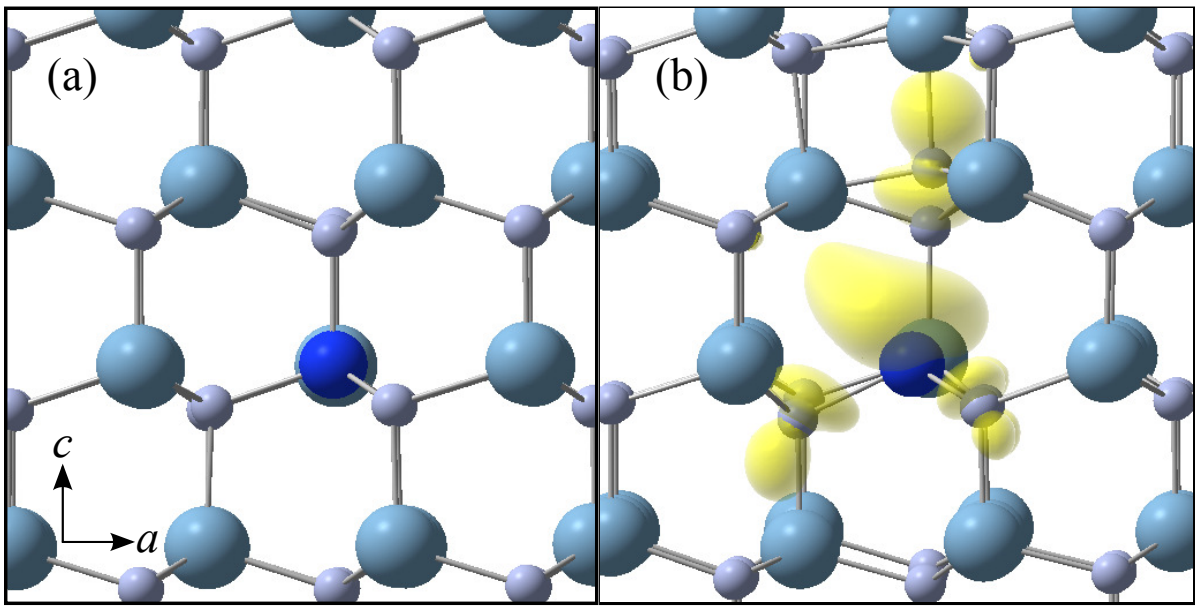

Figure 2.4: Atomic structures of (a) Si^{+}_{AlN} and (b) Si^{-}_{AlN} in the most stable DX-center configuration in AlN. The silicon atom (large dark blue sphere) sits on the substitutional site normally occupied by aluminum (blue spheres); nitrogen atoms are denoted by small light blue spheres. The yellow lobes in (b) indicate the charge density associated with the occupied DX state in the band gap, with the isosurfaces set to 5% of the maximum.

in $Al_xGa_{1-x}N$ (see Table 2.2), i.e., a much higher Al concentration than for the oxygen impurity. For concentrations higher than 94%, the negatively charged Si DX center will be stable for high Fermi levels. However, because the (+/–) level is within 150 meV of the CBM even at the highest Al concentrations, the Si impurity can still be ionized at room temperature. I.e., even though Si nominally acts as a DX center in $Al_xGa_{1-x}N$ with $x \geq 0.94$, it will still give rise to n-type conductivity.

These results explain the apparently conflicting experimental results on Si-doped AlGaN in which formation of DX centers was observed, but n-type conductivity is still obtained. The DX behavior of the Si impurity may still be observed using targeted experiments, particularly at low temperatures. Our results are in good agreement with those of Son *et al.*,[46] who reported DX behavior based on EPR measurements, with the (+/–) level at 0.14 eV below the CBM. Collazo *et al.*[52] measured the electrical characteristics of AlGaN for various Al compositions and found a sharp decrease in carrier concentration and steep increase in donor activation energy for Al content around 80%-90%, again matching well with our results.

2.3.1.3 Germanium in AlN and GaN

Since Si acts as a shallow donor over almost the entire alloy range, one might expect that Ge may have superior characteristics as a donor. Indeed, the atomic size of silicon is not optimal, as evidenced by the inward relaxations of the neighboring N atoms in

the shallow-donor Si_{Al}^{+} state. It might be expected that the larger atomic size of Ge would lead to a better fit and hence make the shallow-donor state even more favorable. However, our explicit calculations on Ge in AlN and GaN do not support this hypothesis. We find similar *DX* configurations as found for Si, but Ge forms a much more stable *DX* center. As in the Si case, the Ge atom stays close to the substitutional site, with an expansion of 5% of the surrounding N atoms; the axial bond along the *c*-axis is broken due to a large displacement of the N atom.

The formation energies of Ge in AlN and GaN are shown in Fig. 2.5. We find that Ge_{Al} impurities form *DX* centers in AlN, with U=−0.92 eV. The (+/–) transition level occurs at 1.05 eV below the CBM, much deeper than for Si in AlN. In GaN, we calculate Ge to be a shallow donor, with Ge_{Ga}^{+} stable for all Fermi levels inside the band gap, and relaxations of the nearest neighbor nitrogen atoms corresponding to a small, 3.5% contraction of the Ga-N bond length; in comparison, the shallow-donor form of Ge in AlN undergoes an even smaller relaxation of the bond lengths of 1% relative to the bulk bond lengths. We are able to stabilize a *DX* distortion for Ge_{Ga} in GaN, but the (+/–) transition occurs at 0.81 eV above the CBM. The small size-induced distortion of the substitutional Ge in GaN indicates that Ge may make a better dopant in GaN and low-Al content AlGaN than Si does; indeed, Fritze *et al.*[53] recently found Ge to be a more effective dopant in GaN relative to Si, causing less stress and leading to better quality material.

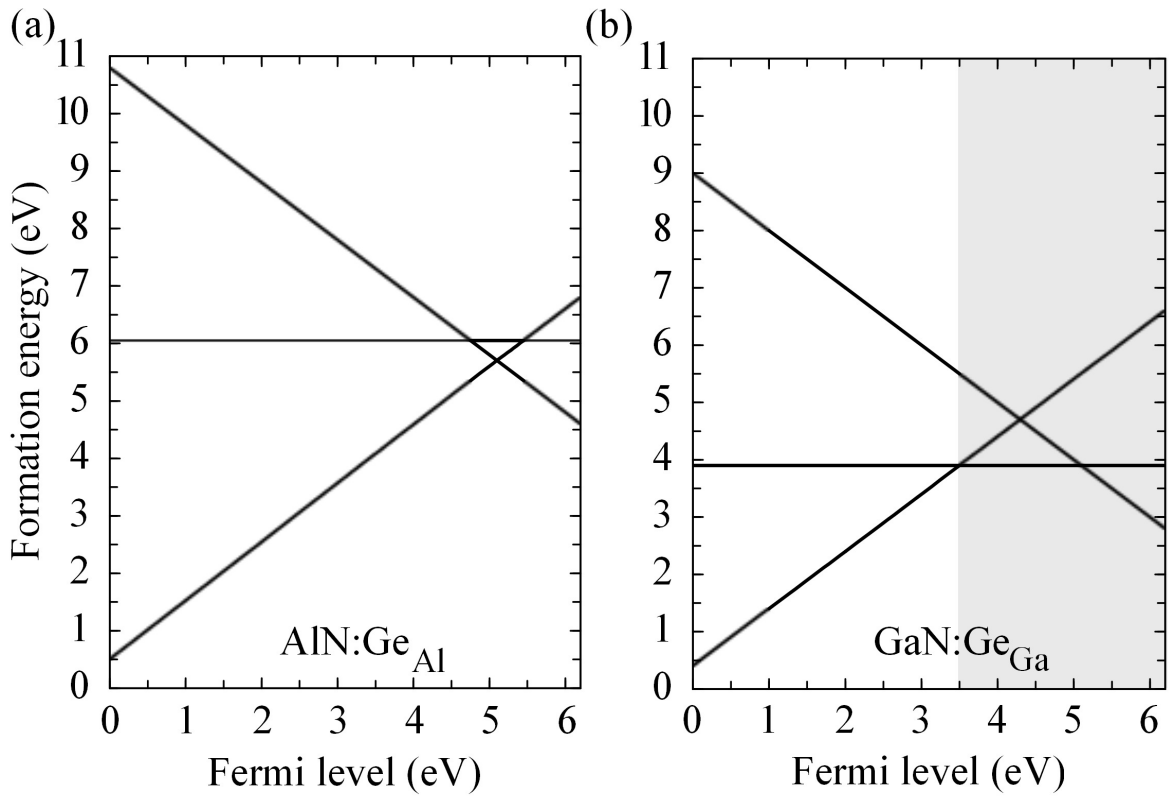

Figure 2.5: Formation energy versus Fermi level for Ge on the cation substitutional site in (a) AlN and (b) GaN, for n-rich conditions. In GaN, Fermi levels above the CBM are indicated by the shaded area.

By interpolating the (+/–) transition levels, and including the effects of band bowing, we find the onset of DX behavior for Ge in $Al_xGa_{1-x}N$ to occur at x=0.52. We are not aware of any attempts to dope AlGaN with Ge, but we predict that the conductivity will drop for Al concentrations higher than about 52% due to the formation of DX centers.

A posteriori we can offer a qualitative explanation for the difference between Si and Ge as donor impurities in AlGaN. The nitride conduction band is composed mainly of Al s states, and Si and Ge when substituted on the Al site result in a s-like impurity state near the CBM. The Ge s state is lower than the Si s state by approximately 1 eV,[54] consistent with the calculated difference in (+/–) levels.

2.3.2 Onset of DX centers in AlGaN and AlInN

While the growth and characterization of pure AlN and GaN is an important subject, for actual devices alloys will typically be used. Specifically, AlGaN and AlInN have applications which range from high-efficiency UV lasers and LEDs[20] to deep UV devices and high-electron-mobility transistors (HEMTs).[55,56] AlGaN has been extensively studied, and there is a large corpus of literature on its growth and electrical properties. In particular, recent experiments by Collazo *et al.*[52] observed a sudden dropoff in carrier concentration for Al contents greater than 80-90%, strongly indicative of DX center formation. By interpolating between the DX transition levels of each system, we can make a prediction for the DX onset alloy concentration in AlGaN, due to O or Si.

By contrast, AlInN has not been explored as much; this is primarily due to difficulties in growth.[57] Specifically, there is a large lattice mismatch between InN and AlN, and there is a large difference in MOCVD growth temperature. However, it can be lattice matched to GaN with a large band gap of around 4.4 eV, which suggests applications in visible and near-UV electronics. In addition, the polarization difference with respect to GaN causes the formation of a high sheet charge density two-dimensional electron gas (2DEG) at the interface. Recent experiments[28] have successfully grown high-quality $Al_{0.82}In_{0.18}N$ alloys by metalorganic chemical vapor deposition (MOCVD). A high level of oxygen incorporation (2 x 10^{18} cm^{-3}) in unintentionally doped AlInN was measured by SIMS, with a donor concentration of 1.1 x 10^{18} cm^{-3} and an acceptor concentration of 6 x 10^{17} cm^{-3}. DX centers were suggested to be the reason for the high acceptor concentration.

In Table 2.2, the +/− levels are shown for oxygen and silicon in InN. One can interpolate between these levels to determine the alloy composition at which the DX transition occurs, listed in the table. It is more difficult to compare with experiment for AlInN than for AlGaN, but we can note that Chung *et al.*[28] observed a high density of compensating acceptors in $Al_{0.82}In_{0.18}N$, which is above the Al concentration predicted from theory. No DX center formation was observed for $Al_{0.82}In_{0.18}N$ doped with Si, which is also in line with our theoretical predictions.

2.4 Effectively n-type doping wide-band-gap alloys

In the previous section, we studied the reasons for decreases in conductivity in III-nitride alloys; in this section, we will specifically study ways in which n-type doping of wide-band-gap nitride alloys can be improved, via alternative dopants, and will suggest sulfur as a very promising alternative to silicon in doping III-nitrides and their alloys. A manuscript containing these results is in preparation.[35]

We have shown that O, Si and Ge all form DX centers in AlN,[27] and the compositions at which shallow-donor-behavior is lost have been predicted. Although Si forms a DX center in AlN, the transition level is within 150 meV of the conduction-band minimum (CBM),[27] and therefore Si can still be used to dope AlGaN and AlN, but not at the levels achieved in InN and GaN.[29] An alternative dopant that would remain a shallow donor throughout the band gap, and avoid DX-center formation altogether, is therefore highly desirable.

The identification of such a new dopant is the goal of this study: we examine potential donors in substitutional configurations, and determine their transition levels and formation energies. Specifically, we examine the behavior of Sn_{Al}, Pb_{Al}, P_{Al}, As_{Al}, Sb_{Al}, S_N, Se_N and F_N. Calculations for Sn_{Al} and Pb_{Al} were performed by my collaborator Joel Varley; all other calculations were performed by the author. Some of the dopants considered, including sulfur, were also studied by Chadi *et al.*[33] However, these calculations

were performed using LDA, and are not reliable; indeed, they predicted that Ge would not form a DX center in AlN, which we have shown to be inaccurate.[27]

2.4.1 A comprehensive study of potential dopants

For the current study, the most relevant quantity is the value of the $(+/0)$ or $(+/-)$ transition level, as this determines the Fermi energy at which the donor is no longer the most favorable charge state. If the transition level level lies within the band gap, it indicates that shallow-donor behavior is suppressed at Fermi energies above that level, and conductivity will be limited.

Formation energy diagrams for Sn_{Al}, Pb_{Al}, P_{Al}, As_{Al}, Sb_{Al}, F_N, S_N and Se_N are shown in Fig. 2.6, under N-rich conditions. In the cases of Sn_{Al}, Pb_{Al}, P_{Al}, As_{Al}, Sb_{Al}, and F_N, the relevant transition level lies well within the band gap, and both Sn and Pb undergo a DX-like transition. This shows that none of these impurities will make good donors in AlN, and will in fact act to inhibit conductivity if present in significant quantities.

However, S_N and Se_N are both shallow donors throughout the band gap, and our results indicate that S_N and Se_N do not undergo the DX transition until the Fermi level is 0.28 eV and 0.41 eV above the CBM, respectively, implying that these defects will act as good shallow donors even in pure AlN. However, we still need to examine whether these impurities are most stable on the N site (where they act as donors), versus incorporating on an Al site or in an interstitial position. To this end, we performed ad-

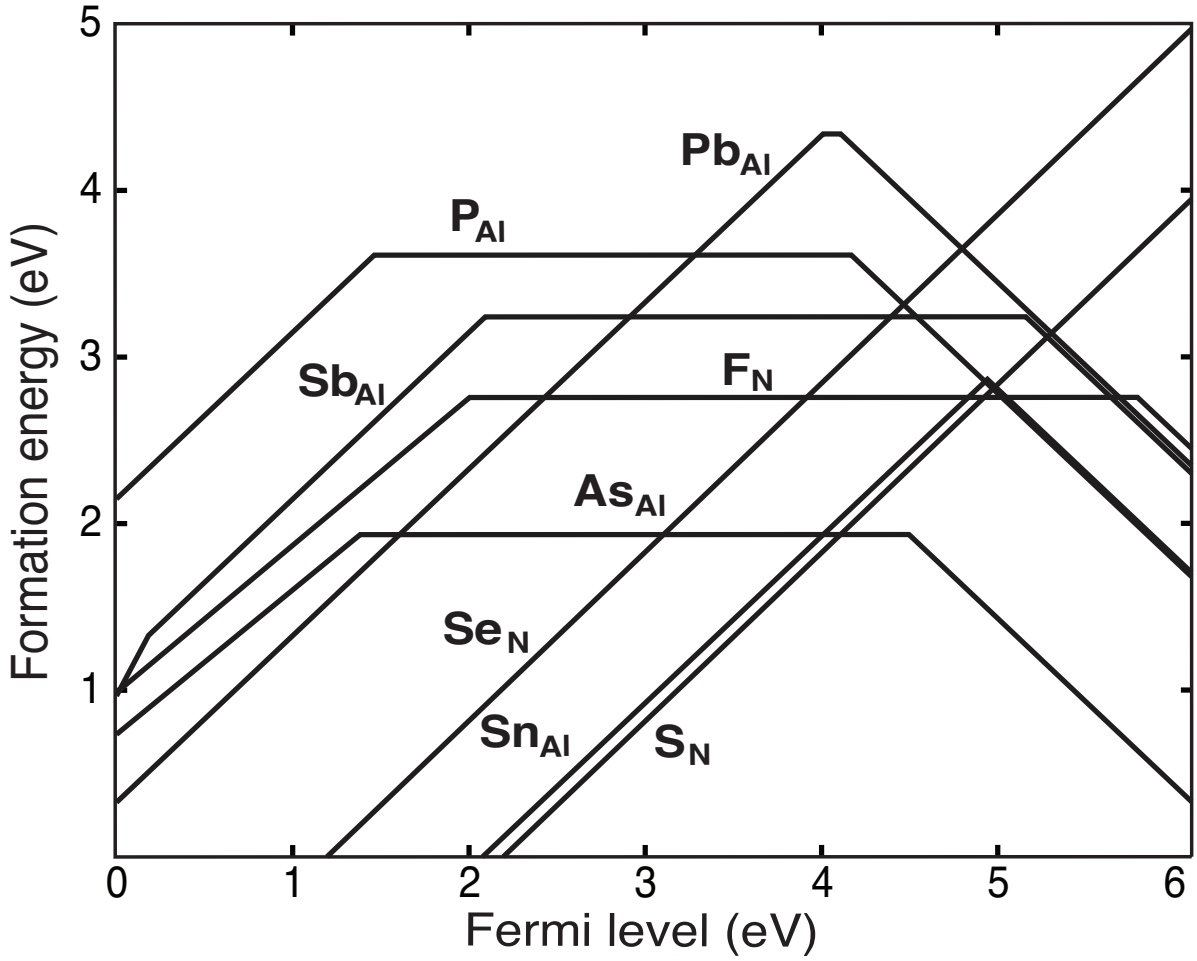

Figure 2.6: Formation energy as a function of Fermi level for Sn_{Al}, Pb_{Al}, P_{Al}, As_{Al}, Sb_{Al}, F_N, S_N and Se_N in AlN, for N-rich conditions.

ditional calculations for these other configurations of S and Se. All interstitial positions tested showed significantly higher formation energies across the band gap than either substitutional site, and can thus be ruled out as highly unlikely to form.

2.4.2 Sulfur and selenium as better n-type dopants

Formation energy diagrams are shown in Fig. 2.7 for Se and Fig. 2.8 for S. Selenium is calculated to be more stable on the Al site for N-rich conditions, but more stable on the N site in Al-rich conditions. The formation energy of the Se_N donor is, however, somewhat high in n-type material: 3.7 eV when E_F is at the CBM, larger than the calculated formation energy of Si under N-rich conditions.[27] Se may be a better donor than Si, but its higher formation energy indicates that it is less likely to incorporate as an n-type dopant in AlN.

Sulfur, on the other hand, is stable on the N site over the entire range of Fermi levels in the band gap, at least when grown under Al-rich conditions. Under those conditions it has a formation energy for E_F at the CBM of 2.1 eV, comparable to the value calculated for Si under n-rich conditions.[27] Combined with the fact that S_N does not undergo a DX transition for any Fermi level within the band gap, we can conclude that S is a highly promising n-type dopant for AlN and high Al-content nitride alloys.

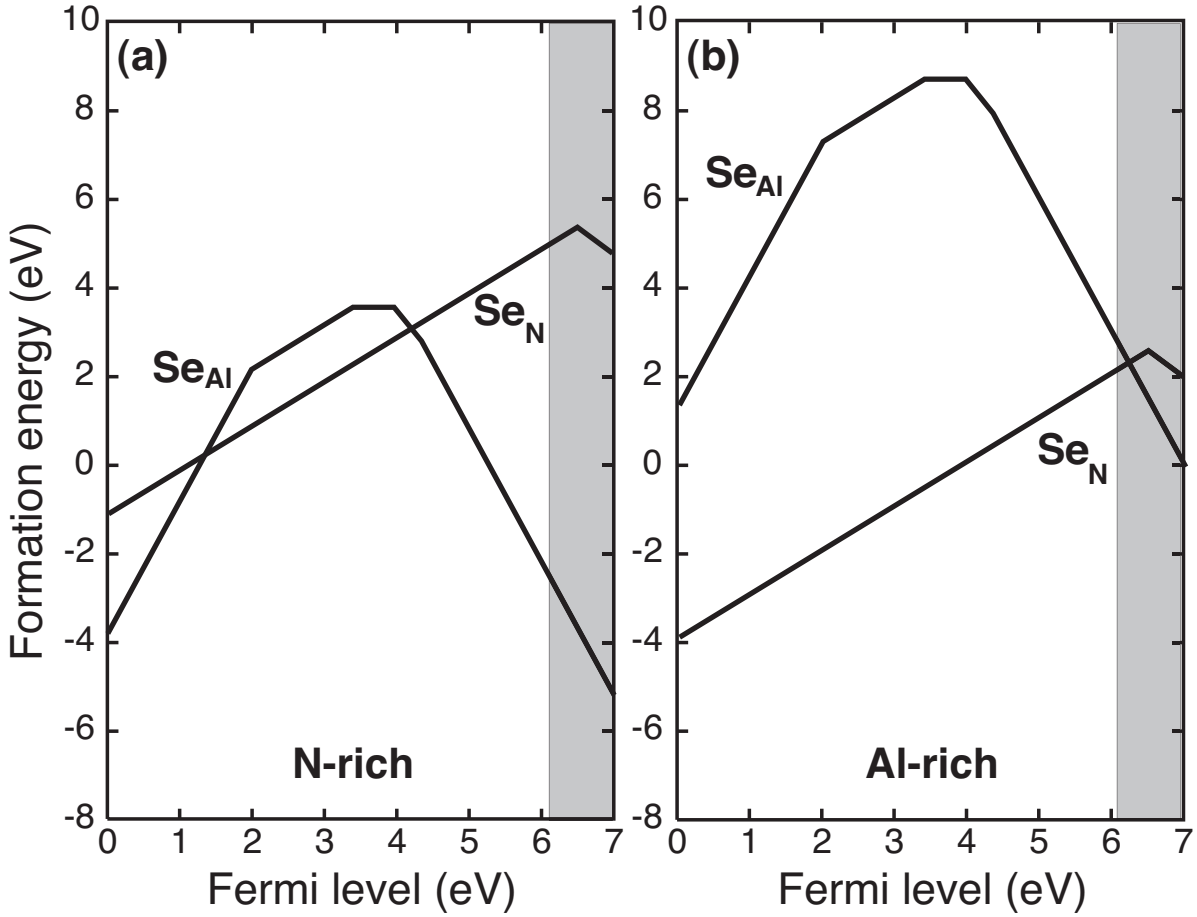

Figure 2.7: Formation energy as a function of Fermi level for Se_{Al} and Se_N in AlN, for (a) n-rich and (b) Al-rich conditions. Fermi levels above the CBM (at 6.1eV) are indicated by the shaded area.

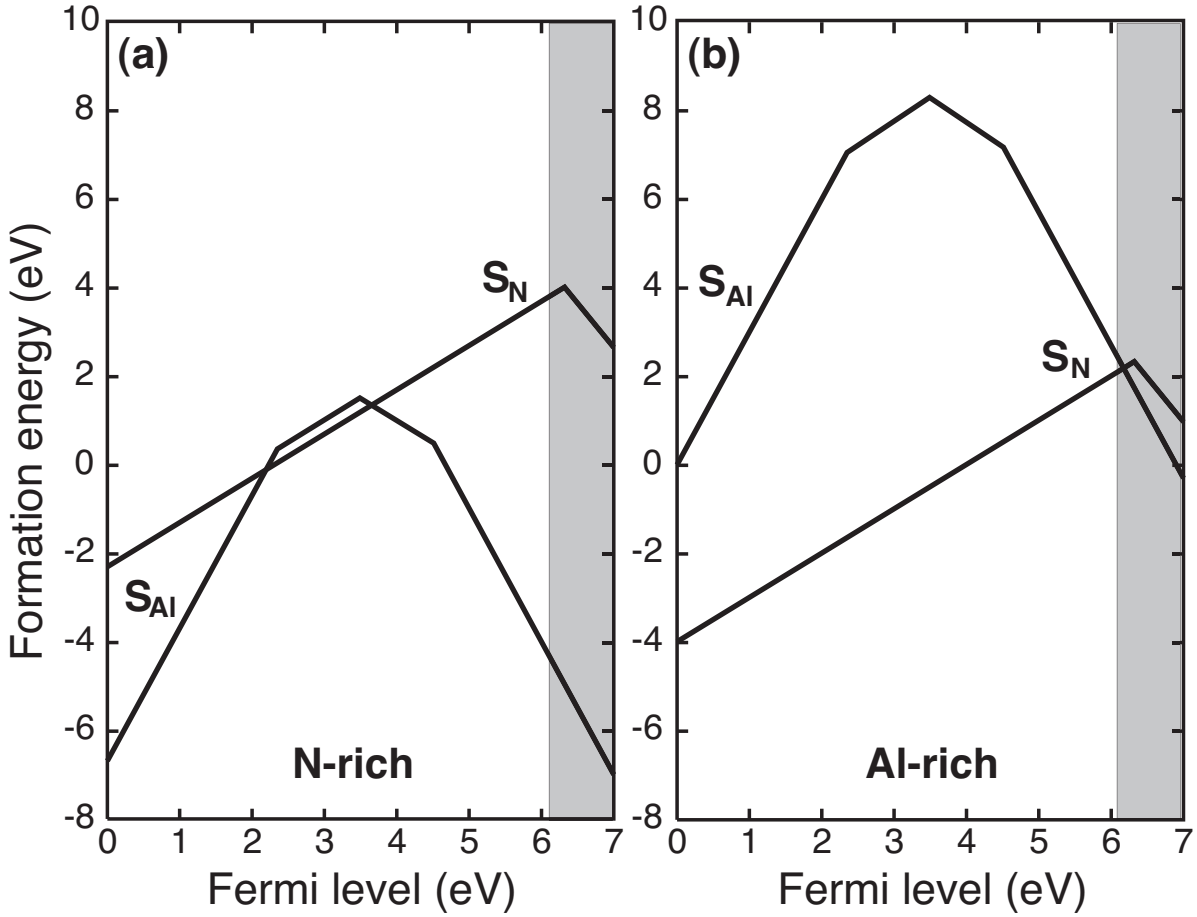

Figure 2.8: Formation energy as a function of Fermi level for S_{Al} and S_N in AlN, for (a) n-rich and (b) Al-rich conditions. Fermi levels above the CBM (at 6.1 eV) are indicated by the shaded area.

2.5 Conclusions

We have studied the role of DX centers O, Si and Ge in AlN and high-Al-content nitride alloys, and have found that all of these impurities form DX centers in AlN, although the Si DX center has a sufficiently low ionization energy to be ionized at room temperature. We have predicted specific Al concentrations for AlGaN and AlInN at which each impurity undergoes a transition from the shallow donor state to the deep acceptor, DX center, state. We have studied a wide range of potential n-type dopants in AlN, and have found that Sn_{Al}, Pb_{Al}, P_{Al}, As_{Al}, Sb_{Al}, and F_N all form deep levels well within the band gap of AlN, and would therefore inhibit conductivity. Promisingly, S_N and Se_N, remain shallow donors at Fermi energies up to the conduction-band minimum. Both impurities are more likely to incorporate on the Al site when grown in N-rich conditions. However, when grown in Al-rich conditions, S and Se both remain shallow donors over the entire band gap range. In particular, S_N has a lower formation energy than Si_{Al}, and does not undergo the DX transition. We propose selenium and especially sulfur as highly promising impurities for future use as n-type dopants in AlN and wide-band-gap III-nitride alloys.

Chapter 3

Defects for quantum computing: a double-edged sword

> At any rate, it seems that the laws of physics present no barrier to reducing the size of computers until bits are the size of atoms, and quantum behavior holds dominant sway.
>
> Richard Feynman

3.1 Introduction

Quantum mechanics has, for many decades, been the basic theory we use to describe the fundamental physics of atoms, and this has allowed us to make predictions regarding phenomena on a much larger scale. From a device perspective, quantum mechanics has long been seen as a problem, rather than a solution. In the binary world of classical bits, the indeterminacy at the heart of quantum mechanics implies a fundamental limit to the size of transistors.

However, starting in the 1980s, a dramatic transformation took place in the understanding of the device implications of quantum mechanics. Rather than being an entirely negative limitation on the size of classically-understood bits, it was realized that the entanglement and non-local phenomena inherent to quantum mechanics could be leveraged into an entirely new kind of device: a quantum computer. A quantum computer could, by utilizing clever techniques such as Shor's algorithm (factoring of large numbers into their component primes) or Grover's algorithm (searching unsorted databases),[12] solve problems that are, for all intents and purposes, completely intractable to classical computers. In addition, a quantum computer could be used to simulate quantum systems exactly, in lieu of our current approximate methods.

The actual development of a functioning quantum computer has remained elusive, despite the vast amount of time and effort put into identifying quantum bits (qubits)

and postulating new types of quantum computer by research groups around the globe. Developing a quantum computer requires complete control over a qubit: i.e., one must be able to initialize the qubit into a defined ground state, manipulate the qubit between two coherent quantum states, and read out the state of the qubit. In addition, a functioning quantum computer would require the entanglement of numerous qubits—for useful factorization using Shor's algorithm, at least 1000 qubits would be required.

In this chapter, we will discuss how atomic-scale defects can act as both a curse for the development of quantum computers by contributing to decoherence at an atomic scale; and as a blessing, by acting as wholly new qubits in new materials. We shall first investigate the origin of dissipative two-level systems in Al_2O_3, and show that hydrogen incorporating in a double-well configuration is a major source of this decoherence (Section 3.2). Next, we will study atomic-scale, paramagnetic defects as potential qubits in new materials, and present results on new defects as qubits in SiC (Section 3.3). All results on H in Al_2O_3 have been compiled into a manuscript,[58] which has recently been submitted for publication. Most of the discussion on defects as qubits has been published in Ref. 59. Manuscripts containing all other results on defects in SiC are in preparation.[60,61]

3.2 Two-level systems in alumina: the problem of defects

The existence of dissipative two-level systems (TLSs), typically observed in amorphous solids, is a long-standing problem in solid-state physics,[62,63] but specific microscopic models have been lacking. A new impetus for uncovering their origins has emerged with the advance of qubits based on superconducting Josephson tunnel junctions.[64,65] Dissipation in the qubit causes the excited $|1>$ state to decay to the $|0>$ ground state, leading to decoherence.[66,67] Two-level systems (TLS) in the insulating layer have been reported to be a major source of energy loss;[68] this mechanism seems to dominate other sources of decoherence.

Dielectric loss from TLS can be large in amorphous materials (see Pohl *et al.*[69] for a detailed survey), and is thought to arise from random bonding of atoms. In superconducting qubits, TLSs are a source of decoherence in the tunnel barrier of the Josephson junction. Loss arises from absorption of microwave radiation by TLSs with an electric dipole moment. It can be modeled as atoms tunneling between two distinct positions,[68] and has been shown to be important even for the surface oxide of superconducting metals (Al_2O_3).[70] The effect of decoherence can be mitigated by the use of single-crystal Al_2O_3; a reduction in the density of spectral splittings of up to 80% has been observed.[71] The defects have resonance frequencies on the order of 10 GHz,[68] comparable to the qubit

circuit; the coupling strengths and decoherence times are sufficiently large for coherent oscillations between the qubit and TLS. Despite extensive studies on the physics and effects of TLSs,[62,63,72] their microscopic origin has remained unsettled.

Here we show that TLSs in Al_2O_3 can be attributed to hydrogen impurities that incorporate on interstitial sites. Hydrogen is an ubiquitous impurity, present in many growth and processing environments, and able to unintentionally incorporate in most materials,[73] including the Al_2O_3 dielectric used in superconducting qubits. The specific characteristics of hydrogen that give rise to the TLS are related to its propensity for hydrogen bonding; by definition it is the only element that exhibits this type of chemical bond. Interstitial hydrogen in oxides exhibits a strong, mainly covalent bond with a primary O atom, with a typical O-H bond length of $\sim$1 Å, but can also interact with a next-nearest-neighbor (*nnn*) O atom, resulting in an O-H...O configuration.[74,75] For suitable O-O distances, the interaction with the *nnn* O atom leads to quantum-mechanical tunneling between these neighboring sites. Tunneling of interstitial hydrogen between adjacent O sites has been observed in oxides such as $KTaO_3$, for which Spahr *et al.*[76] reported tunneling rates in the 7-40 GHz range. We will demonstrate, based on first-principles theory, that hydrogen in Al_2O_3 can give rise to TLSs with tunneling frequencies that explain the dielectric loss observed in superconducting qubits.

3.2.1 Methodology: a novel approach

Our calculations are based on hybrid density functional theory (DFT-HSE), as described in Section 1.2.3. We use a Hartree-Fock mixing parameter of 32%, chosen to correctly describe the band gap of α-Al_2O_3.[77] The impurity calculations were performed using a supercell containing 120 atoms; this is a 2×2×1 multiplier of the 30-atom unit cell of α-Al_2O_3. We used a 2×2×1 Monkhorst-Pack **k**-point mesh for the integrations over the Brillouin zone, and a cutoff of 500 eV for the plane-wave basis set. Our HSE calculations for α-Al_2O_3 produce lattice parameters a=b=4.74 Åand c=12.95 Å, in very good agreement with the experimental values, a=b=4.76 Å and c=12.99 Å.[78]

We study a variety of O-O interatomic distances (representative of those occurring in the amorphous phase or near surfaces or interfaces) by varying the volume of the crystalline solid, up to a total volumetric compression of 6%. The physics of the TLS studied here is determined by the local environment of the hydrogen atom, for which the presence (or absence) of long-range order is irrelevant. Our results for crystalline α-Al_2O_3 therefore also apply to amorphous Al_2O_3, in which the relevant O-O interatomic distances occur, as evidenced by pair-correlation distributions.[79,80]

The 3D potential energy surface for the H atom in the host lattice (fixed to the coincidence configuration) was determined by calculating the total energy for each point on a grid with a spacing of 0.24 Å in the three spatial directions. This grid was then

interpolated using the energy gradients obtained from the forces acting on the H atom, leading to a smooth function. To numerically solve the Schrödinger's equation for the quantum-mechanical motion of H, the wave functions were expanded in a plane-wave basis set with a cutoff energy of 800 eV; this cutoff was determined based on convergence tests for the tunnel splittings.

The qubit-TLS coupling S_{Max} can be calculated as[68,81]

$$S_{Max} = \frac{2p}{x}\sqrt{e^2 E_{01}}2C, \tag{3.1}$$

where x is the barrier thickness, p is the effective dipole moment of the charge in the double-well system, C is the capacitance, and E_{01} is the qubit splitting energy. We assume a junction width $x = 2$ nm, a capacitance $C = 850$ fF, and a qubit splitting energy 5.4 GHz, corresponding to representative experimental values.[65,82] The dipole moment is calculated as the product of the effective charge around the hydrogen atom in the coincidence configuration and the distance between the symmetric potential wells, determined from first-principles calculations.

3.2.2 Calculations on hydrogen in Al_2O_3

The process of quantum-mechanical tunneling of interstitial hydrogen is complicated by the strong interactions of hydrogen with the lattice, which lead to relaxations of the host.[83] In its most stable configuration, the interstitial H is primarily bonded to one O

atom, with a larger distance to the *nnn* O (Fig. 3.1(a)). If the positions of the host atoms are fixed to that configuration, the potential energy surface for the H atom is highly asymmetric, i.e., the potential well near the *nnn* O is high in energy (Fig. 3.1(c)), or even nonexistent. To enable tunneling, the lattice must be brought into a symmetric *coincidence configuration*[83,84] (Fig. 3.1(b)) where a double-well potential occurs and tunneling can take place (Fig. 3.1(c)). The three-dimensional (3D) potential energy surface corresponding to the coincidence configuration determines the energy levels and tunnel splittings for the quantum-mechanical motion of the H atom. These energy levels and tunneling frequencies can be calculated by numerically solving Schrödinger's equation. The determination of the coincidence configuration and the corresponding potential energy surface is an important aspect of the present work, discussed in detail below. We note that considering a purely adiabatic process, in which the host atoms assume their lowest-energy positions for each location of the H atom, is unrealistic due to the large mass difference between H and the other atoms. On the other hand, fixing the host atoms to positions corresponding to an unrelaxed lattice, or to those for the ground-state configuration of interstitial H, is unrealistic as well.

Interstitial H in α-Al_2O_3 can assume three different charge states: positive (H_i^+), neutral (H_i^0), and negative (H_i^-). The use of a hybrid functional is absolutely essential for properly describing the interaction between H and the *nnn* O atom, and thus the tunneling frequencies of the associated TLS. DFT within the standard generalized

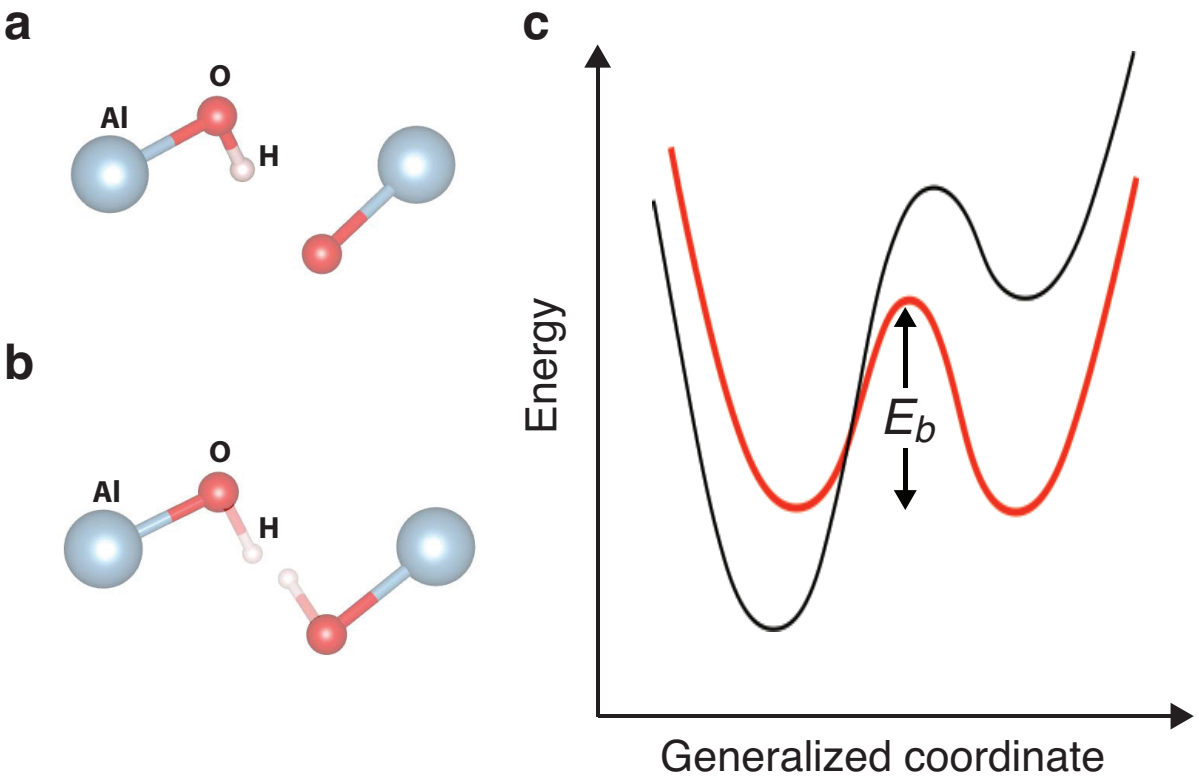

Figure 3.1: Self-trapped and coincidence configurations. (a) Geometry of the self-trapped configuration for a hydrogen interstitial in Al_2O_3. H is bonded primarily to a specific O atom, but could alternatively be bonded to a *nnn* O atom. Al atoms are represented by large (grey) spheres, O atoms by smaller (red) spheres, and H atoms by the smallest (pink) spheres. (b) Geometry of a coincidence configuration, obtained by averaging over two adjacent self-trapped configurations. The two symmetric hydrogen sites are indicated by semi-transparent bonds. (c) Schematic potential-energy curve for the self-trapped (black) versus the coincidence (red) configurations. The latter corresponds to a double-well system with energy barrier E_b.

gradient approximation (GGA) overestimates the strength of the interaction between H and the *nnn* O atom so that the calculated O-H frequency is too low.[85] For the neutral and negative charge states, the interstitial hydrogen occupies a position in between two Al atoms, as shown in Fig. 3.2. H_i^+, on the other hand, bonds to an O atom with a bond length of 1.01 Å, and a distance of 1.70 Å to the *nnn* O (Fig. 3.1(a), 3.2). The relative stability of the different charge states depends on the Fermi-level position. The formation energy of an interstitial H in charge state q [$E^f(H_i^q)$] is calculated as:[7]

$$E^f(H_i^q) = E_{\rm tot}(H_i^q) - E_{\rm tot}(Al_2O_3) - \mu_{\rm H} + q\epsilon_F, \tag{3.2}$$

where $E_{\rm tot}(H_i^q)$ is the total energy of the supercell containing an interstitial H atom in charge state q, $E_{\rm tot}(Al_2O_3)$ is the total energy of the bulk supercell, $\mu_{\rm H}$ is the hydrogen chemical potential, and ϵ_F is the Fermi level, referenced to the valence-band maximum (VBM). The value of μ_H does not affect the relative stability of the different configurations, and we can set μ_H equal to half the total energy of an isolated H_2 molecule. The calculated formation energies as a function of ϵ_F are shown in Fig. 3.3. We find that the donor state, H_i^+, is the stable charge state for Fermi levels up to 5.9 eV above the VBM; above that, the acceptor state, H_i^-, is most stable. The neutral charge state, H_i^0, is always higher in energy than both the H_i^+ and H_i^-, reflecting a negative-U character, as observed for H_i in many other semiconductors and insulators.[73] The position of the Fermi level is determined by charge neutrality; in an undoped α-Al_2O_3 crystal, native

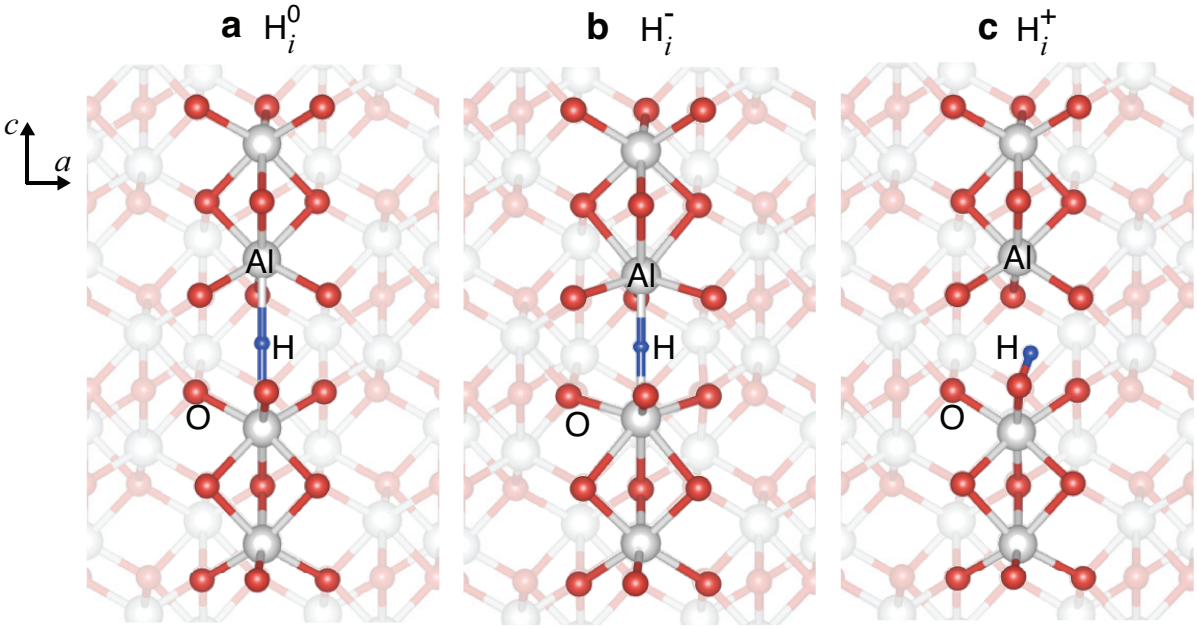

Figure 3.2: Geometry of self-trapped hydrogen configuration in Al_2O_3. Relaxed "self-trapped" geometry of interstitial H in the (a) neutral, (b) negative, and (c) positive charge states. Al atoms are represented by large (grey) spheres, O atoms by smaller (red) spheres, and H atoms by the smallest (blue) spheres.

defects lead to a Fermi-level position around the middle of the gap.[86] Under these conditions, Fig. 3.3 shows that interstitial hydrogen is most stable in the positive charge state, H_i^+.

The relaxed geometry of the H_i^+ configuration in Al_2O_3 at its equilibrium volume is asymmetric, with significant relaxations of the host atoms, particularly of the O atom to which it is bonded (Fig. 3.2(a)). The energy difference between the unrelaxed Al_2O_3 lattice containing an H_i^+ and the relaxed lattice with H_i^+ in its most stable configuration is defined as the self-trapping energy,[84,87] and is approximately 1.5 eV. One would expect this large energy difference to prohibit H from tunneling. In order to enable tunneling, a configuration must be created in which the hydrogen is equally likely to be bonded to either of the two neighboring O atoms. The lowest energy required to take a self-trapped configuration into a symmetric structure, i.e., the coincidence configuration, is the "coincidence energy" E_c. The formation of a coincidence geometry may be assisted by lattice vibrations, or may occur in regions where the atomic arrangement deviates from that in the bulk crystal, such as near surfaces or interfaces, or in an amorphous phase, in which a range of O-O distances from around 2.4 Å to 2.8 Å are observed.[79,80]

An exact determination of the coincidence geometry would in principle require a self-consistent treatment of the quantum-mechanical hydrogen motion coupled to the host-atom relaxation, something that is too computationally demanding to be performed in conjunction with a first-principles treatment of the electronic structure. Approximate

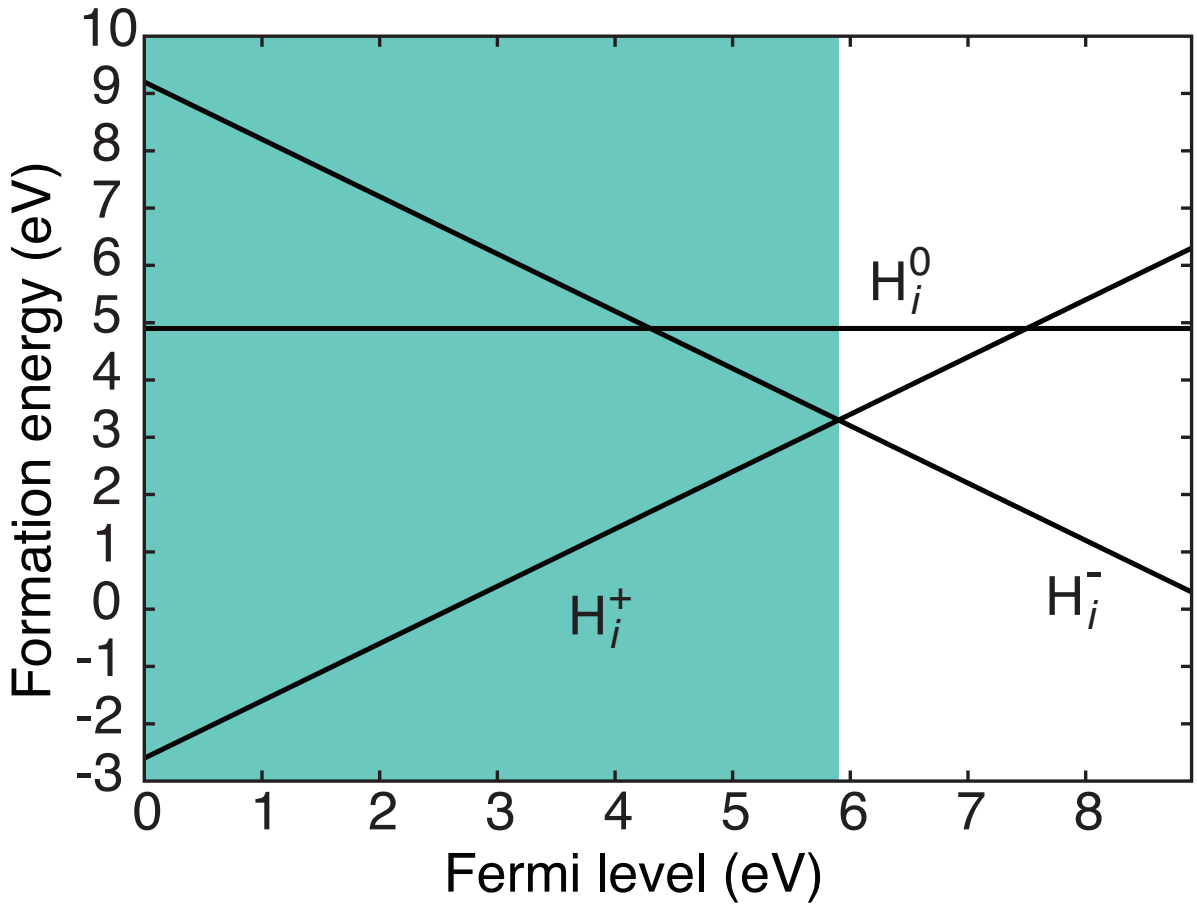

Figure 3.3: Formation energy of H in Al_2O_3. Formation energy, calculated with DFT-HSE, of various charge states of interstitial hydrogen in α-Al_2O_3 as a function of Fermi level. The shaded area indicates the region of Fermi-level values in which H forms a hydrogen bond.

methodologies for obtaining the coincidence geometry have therefore been developed. Here we determine the coincidence configuration by averaging the positions of the host atoms for two self-trapped configurations, corresponding to H bonded to adjacent O sites. This procedure produces a symmetric double well potential for the tunneling hydrogen atom. We then place a H atom (H_i^+) in the structure formed by the host atoms in the coincidence geometry and allow the hydrogen to relax (keeping the host atoms fixed). The energy difference between the relaxed H_i^+ in the coincidence geometry and the H_i^+ in the most stable, self-trapped configuration is the coincidence energy E_c.

3.2.3 Hydrogen as the origin of dielectric loss in superconducting qubits

We determined the coincidence configurations and coincidence energies for a range of different volumes of α-Al_2O_3, from equilibrium to 6% isotropic (linear) strain. This variation in lattice parameters corresponds to O-O distances between 2.715 to 2.59 Å in the perfect crystal, a range which includes the O-O distances found in amorphous Al_2O_3.[79,80] The corresponding O-O distance in the coincidence configuration varies in the range of 2.56 to 2.45 Å. The results are listed in Table 3.1 and shown in Fig. 3.4. The coincidence energies vary from 0.36 to 0.25 eV, with a minimum at 4% compression, corresponding to an O-O distance (in the coincidence geometry) of 2.49 Å.

Table 3.1: Calculated TLS parameters for H in Al_2O_3. Coincidence energies E_c, tunnel splittings Δ, tunneling frequencies ν, and coupling constants S_{Max} for two-level systems associated with interstitial hydrogen in α-Al_2O_3. Values are listed for various O-O distances d_{O-O} ranging from 2.56 Å to 2.45 Å, corresponding to volume compression by the specified amounts.

Compression	d_{O-O} (Å)	E_c (eV)	Δ (eV)	ν (GHz)	S_{Max} (MHz)
0%	2.56	0.36	2.20 x 10^{-4}	53	19.8
1%	2.55	0.35	2.09 x 10^{-4}	50	18.3
2%	2.53	0.33	2.48 x 10^{-4}	59	19.2
3%	2.51	0.32	2.17 x 10^{-4}	53	17.1
3.5%	2.50	0.29	1.38 x 10^{-4}	33	16.7
4%	2.49	0.25	1.07 x 10^{-4}	26	16.4
4.5%	2.48	0.28	6.7 x 10^{-5}	16	16.4
5%	2.47	0.28	4.12 x 10^{-4}	90	16.3
6%	2.45	0.26	1.57 x 10^{-3}	360	16.0

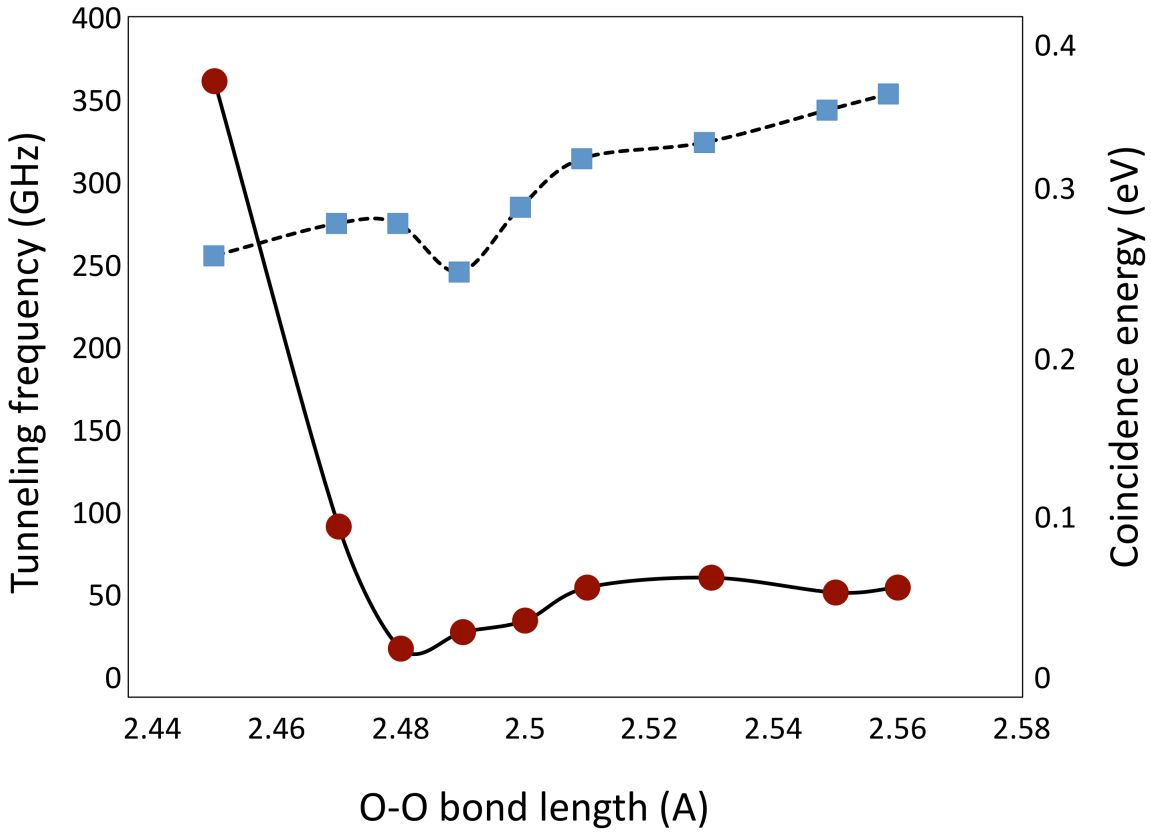

Figure 3.4: Tunneling frequency as a function of O-O distance. Calculated tunneling frequencies and coincidence energies are shown for O-O distances ranging from 2.45 to 2.56 Åin the coincidence geometry. Tunneling frequencies are shown in red, coincidence energies in blue.

The solution of Schrödinger's equation for the H atom in the 3D potential energy surface corresponding to the coincidence geometry produces wave functions and energies characteristic of a double-well potential. The tunnel splitting is calculated from the energy difference between the ground state and the first excited state, and leads to the tunneling frequencies listed in Table 3.1. In order to enable direct comparison to experiments, we have also calculated the qubit-TLS coupling strength for each possible configuration. The measured signal of a TLS in a phase qubit reflects the resonance between the splitting energy E_{01} and the tunnel splitting of the TLS, expressed as the qubit-TLS coupling strength S_{Max}. As described in the Section 3.2.1, S_{Max} is a function of the splitting E_{01} and the dipole moment p. We have calculated this quantity using experimental values for the qubit frequency (5.4GHz,[82]) and for the capacitance and gate width, resulting in values between 16 and 20 MHz (Table 3.1).

The results in Table 3.1 indicate that the tunneling frequencies are sensitive to the O-O distance, albeit not in a straightforward manner. Compression of the lattice initially makes little difference to the tunneling frequency, which remains around the value of 50 GHz calculated for the equilibrium volume (see Fig. 3.4). The higher values of E_c (compared to those calculated at compressed volumes) suppress the possibility of tunneling in this regime. At 4.5% compression, corresponding to an O-O distance of 2.59 Å in the impurity-free crystal, the O-O distance in the coincidence configuration is 2.48 Å and the tunneling frequency is 16 GHz. Further compression of the lattice leads

to a decrease in the potential barrier between the wells in the double well potential, and thus an increase in the tunneling frequency, away from the range observed in experiment. Such small O-O distances also have a decreased probability of occurring in actual samples.

Experimentally, a range of TLS tunneling frequencies have been reported, centered around 10 GHz.[65,68] To further corroborate our model, we can also compare our calculated qubit-TLS coupling constants S_{Max} to experiment. Reported values range from 25 MHz to 45 MHz.[65,88] The agreement with our calculated values (16-20 MHz, Table 3.1) further validates our claim that hydrogen is responsible for TLS decoherence in the dielectric layer.

The presence of hydrogen in dielectrics is highly plausible, particularly when they are deposited or grown using hydrogen-containing precursors. Hydrogen has been observed with secondary ion mass spectrometry (SIMS) in atomic-layer-deposited oxides.[89] Unfortunately, accurate experimental detection of hydrogen is fraught with difficulty, at least at the low concentrations and in the small volumes that are relevant for qubits. A decrease in loss has been reported in the case of crystalline dielectrics.[71] This is consistent with our results for crystalline α-Al_2O_3 at equilibrium volume, in which the tunneling frequencies are well outside the range that is relevant for the qubits. However, even in devices that use crystalline dielectrics, there is still a range of O-O distances at surfaces and interfaces,[90] so TLS losses are never completely eliminated.

The results reported here are distinct from recent work by Holder *et al.*,[72] also based on DFT-HSE, which considered the rotor motion of a H atom in α-Al_2O_3, either as an interstitial or in the Al vacancy. They concluded that interstitial hydrogen could not be responsible for TLSs, based on a value for the calculated tunneling frequency that was much higher than experiment. They did find one defect displaying frequencies in the GHz range, namely a complex of H with an Al vacancy in the +1 charge state. However, this complex is stable only if the Fermi level is within 1.2 eV of the VBM, which is highly unlikely (essentially impossible) in α-Al_2O_3, a wide-band-gap insulator. The frequency calculated for H_i^+ in Ref. 72 ($\sim$240 GHz) is significantly higher than the values reported here, likely stemming from their highly simplified 1D model. Our own tests have indicated that a full mapping of the 3D potential energy surface is essential in Al_2O_3, and that 1D models do not match the full results. In addition, the sixfold or threefold degeneracies assumed in the rotor model of Ref. 72 are extremely unlikely to occur in real materials—even in crystalline solids. In contrast, the two-fold degeneracy inherent in our hydrogen-related double-well systems is highly plausible and fits within the general framework of two-level systems.

In summary, we have proposed a microscopic model for TLSs in Al_2O_3 based on interstitial hydrogen in an O-H...O hydrogen-bonding configuration. Hybrid DFT calculations produce frequencies that are in the range reported for TLSs known to be responsible for the main loss in superconducting qubits, and calculated qubit-TLS cou-

pling paramaters are close to experimental values. We suggest that hydrogen could be responsible for TLSs in other materials as well, given its ubiquity. As the only element to exhibit "hydrogen bonding", it stands out as a candidate for TLSs in oxides, in which a range of suitable O-O distances occur. The low migration barriers observed (and calculated) for interstitial H^+ suggest that barriers between equivalent sites can be low enough to lead to significant tunneling and hence double-well systems.

3.3 New atomic-scale qubits in materials: the solution of defects

A qubit is the basic unit of information in a quantum computer. In contrast to binary bits that can have only two values (0 or 1) and upon which Boolean logic is based, a qubit can be in any coherent superposition between two quantum states. The physical realization and control of qubits is very challenging. A range of qubit implementations are being explored, including in liquids, atoms, and in solids such as superconductors and semiconductors. Loss of coherence is a major issue: in principle, the qubit should be completely isolated from unwanted external fluctuations in its environment in order to maintain its quantum state. However, a completely isolated qubit would not allow for the interactions, such as entanglement, that are necessary to actually perform quantum

manipulations and computing. Even if a single qubit can be fabricated, scalability is a major issue.

The successful development of quantum computers is dependent on identifying quantum systems to function as qubits. It has recently been found that certain point defects in wide-band-gap semiconductors can serve as qubits for quantum computing.[91] These centers are stable in well defined spin states that can be initialized and manipulated at room temperature using optical and microwave excitations, and are responsive to electric and magnetic fields. Paramagnetic states of point defects in semiconductors or insulators have been shown to provide an effective implementation, the nitrogen-vacancy center in diamond being a prominent example. The spin-one ground state of this center can be initialized, manipulated, and read out at room temperature. Identifying defects with similar properties in other materials would add flexibility in device design and possibly lead to superior performance or greater functionality. A systematic search for defect-based qubits has been initiated, starting from a list of physical criteria that such centers and their hosts should satisfy.[92] First-principles calculations of atomic and electronic structure are essential in supporting this quest: they provide a deeper understanding of defects that are already being exploited and allow efficient exploration of new materials systems and defects by design.

In principle, a wave function on an isolated atom would provide an intuitive, well-defined, and well-understood quantum state for use as a qubit. Unfortunately, isolated

atoms do not easily lend themselves to incorporation in a quantum device; complex approaches such as ion traps or optical lattices are required to constrain the atoms or ions. It turns out, however, that point defects in semiconductors or insulators can display behavior that is very similar to that of isolated atoms, as illustrated in Fig. 3.5. The electronic states associated with a vacancy, for instance, tend to have energies that lie within the forbidden band gap, and their wave functions are constructed out of atomic orbitals on the neighboring atoms and are hence very localizedon the scale of atomic dimensions. Invoking similarities with wave functions on isolated atoms is therefore quite appropriate. The advantage of point defects is, of course, that they are firmly embedded within the host material, and decades of investigation and characterization have provided us with many tools for controlling and manipulating such defects.

Point defects tend to have a bad reputation. When unintentionally present in semiconductors, they can adversely affect the desired doping behavior and lead to degraded electronic or optical properties. The semiconductor community has therefore gone to great lengths to build a thorough understanding of such defects and to develop exquisitely precise characterization techniques, such as electron spin resonance and photoluminescence. This knowledge and tool set can now be constructively employed to design and manipulate point defects for use as qubits.

The schematic provided in Fig. 3.5 is a qualitatively correct but obviously very simplistic description. A serious approach to designing defects needs to be based on

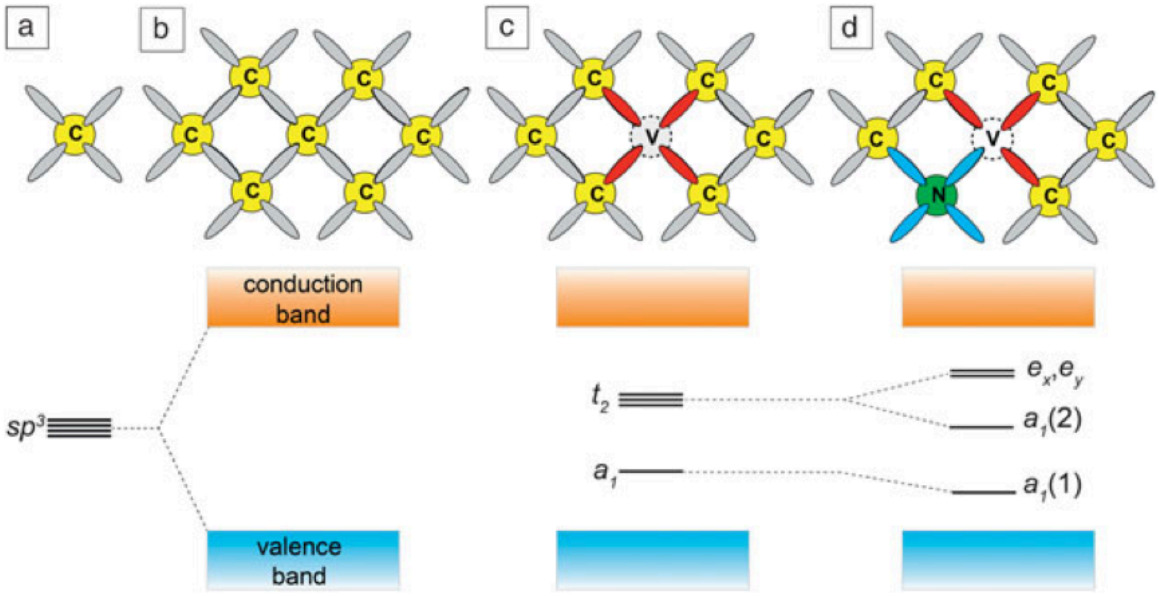

Figure 3.5: Schematic representation of the electronic structure of a point defect in a tetrahedrally coordinated elemental semiconductor such as diamond. (a) The electronic states corresponding to the sp^3 orbitals on an isolated C atom. (b) The superposition of these orbitals that gives rise to the band structure of an infinite solid. If a carbon atom is removed, as shown in (c), a vacancy is created, and the four orbitals on the surrounding atoms interact with each other in the tetrahedral environment to give rise to states with a_1 and t_2 symmetry. A symmetry-lowering perturbation, such as incorporation of a nitrogen atom on one of the sites around the vacancy (d), further splits the t_2 states.

a rigorous quantum-mechanical description. For isolated atoms, we would solve the Schrödinger equation. The electronic structure of a solid, however, presents a very complicated many-body problem.

The prototype of this class of qubits is the NV center in diamond, which is composed of a substitutional N impurity next to a vacancy (Fig. 3.5(d)). In the negative charge state [$(NV)^-$], optical transitions between the S=1 ground state ($^3A_2$) and the ^3E excited state result in luminescence that can be modulated by an external magnetic field.[93] It has also been demonstrated that the $(NV)^-$ center exhibits long spin-coherence times at room temperature.[94] Fig. 3.6 illustrates the electronic structure of this center, as calculated using hybrid DFT.

The positioning of the defect states within the gap follows the generic pattern of Fig. 3.5(d) (with the addition of spin polarization), but quantitative differences occur. For instance, the $a_1(1)$ states are not located within the band gap but are resonant in the valence band. These bonding states do not play a direct role in the functionality as a spin center, and hence their energy is not directly relevant. The states that are of key importance are the spin-down $a_1(2)$ and e_x, e_y states: laser light of appropriate wavelength can cause a spin-conserving transition in which a spin-down electron is excited from an $a_1(2)$ state to an e_x/e_y state; this corresponds to excitation from an $^3A_2$ ground state to an ^3E excited state. The details of the operation of the NV center require a description in terms of such multiparticle states. However, the single-particle

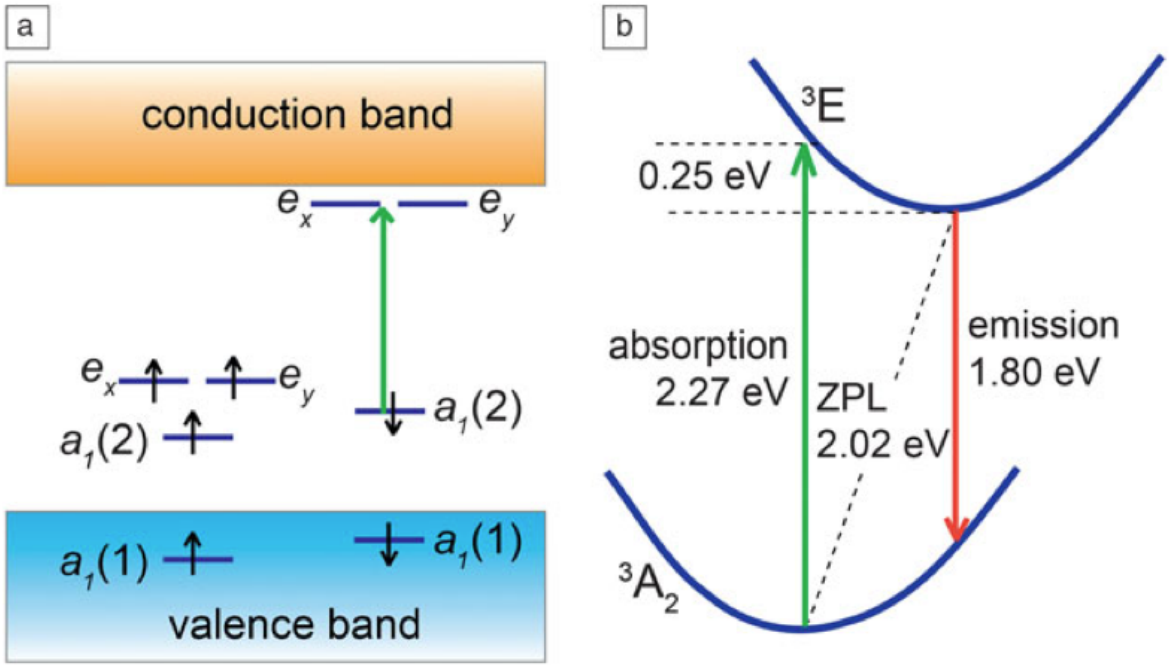

Figure 3.6: (a) Electronic structure of the negatively charged nitrogen-vacancy (NV) center (NV) in diamond, as calculated with DFT-HSE by my collaborators Justin Weber and Joel Varley in Ref. 92. Optical excitation (vertical green arrow) can lift an electron out of the spin-down $a_1(2)$ state into an e_x/e_y state. (b) Calculated configuration coordinate diagram for the NV center. The lower curve indicates the energy of the defect in its electronic ground-state configuration ($^3A_2$) as a function of a generalized coordinate, which measures the displacements of atoms. The upper curve corresponds to the ^{3}E excited state. The zero-phonon line (ZPL) represents a transition between the two configurations in their relaxed atomic configurations; the intensity of this ZPL tends to be weak if these atomic configurations are very different (i.e., if large relaxations occur). Peaks in the optical absorption and emission curves will correspond to the vertical transitions (green and red arrows) for which the atomic positions remain fixed.

picture presented in Fig. 3.6 is sufficient to discuss the basic features of the NV center that are key to designing similar point-defect centers. The defect states follow the pattern of Fig. 3.5(d), but spin polarization leads to different energies for spin-up and spin-down states. Six electrons need to be accommodated in the defect states: each dangling bond on a C atom contributes one electron, while the dangling bond on the N atom contributes two electrons due to the higher valence of nitrogen. In addition, an extra electron (provided by donors elsewhere in the material) is present that puts the center in a negative charge state. Filling the electronic states in order of increasing energy leads to the occupation shown in the figure, resulting in a $S = 1$ (triplet) state for the center.

The optical excitation energy corresponding to the transition depicted by the arrow in Fig. 3.6(a) is evaluated by constraining the occupation to that of the excited state while keeping the atomic positions fixed to those of the ground state. The calculated energy difference (2.27 eV) is expected to yield the peak energy of the absorption spectrum, since electronic transitions can happen on a time scale much faster than atomic relaxations. If we subsequently allow the atomic positions to relax, maintaining the excited-state triplet electronic configuration, we obtain a relaxation energy of 0.26 eV (the Frank—Condon shift), as illustrated in the diagram of Fig. 3.6(b). This configuration coordinate diagram allows the assessment of key features of optical absorption and emission processes for a point defect in which there may be significant differences in the atomic configurations of

ground state and excited state. The energy difference between the excited state and the ground state, both with relaxed atomic configurations, corresponds to the zero-phonon line, calculated here to be 2.02 eV, and compares well with experiment, which reported a value of 1.945 eV.[95] We can also obtain the peak emission energy, at 1.80 eV, in good agreement with the experimental value of 1.76 eV.

The good agreement with experiment (to within 0.1 eV) is particularly impressive given that the calculations are completely *ab initio* and involved no fitting to any experimental quantities. In addition to optical emission, a non-spin-conserving decay path also exists that includes a nonradiative transition to an intermediate spin-singlet $^1A_1$ state. This transition plays a key role since it allows the center to be optically initialized into a specific spin sublevel. Recent calculations and experiments place the $^1A_1$ state 1.19 eV above the $^3A_2$ ground state.

3.3.1 Identifying potential defects as qubits

The thorough and quantitative understanding achieved for the NV center in diamond renders it an ideal prototype for generating a list of design criteria that other qubit candidates should satisfy, enumerated first by my collaborators Justin Weber and Joel Varley in Ref. 92. The highly localized nature of the bound states of the NV center is critical in making a robust qubit and isolating the center from sources of decoherence; indeed, coherence times of over half a second have recently been demonstrated. In

combination with on-chip microwave-frequency waveguides that enable quantum-control operations on sub-nanosecond time scales, tens of millions of coherent operations can be performed within this spin coherence time.

The localized nature of the NV-center wave functions is directly related to the position of the relevant states within the band gap: states that are close to, or resonant with, the conduction or valence bands of the host will more strongly interact with the extended states and be more delocalized. It is the energetic separation of band-to-band and defect-to-band from intra-defect optical transitions that allows the defect to be initialized and measured with high fidelity at room temperature and even well above room temperature. This offers significant advantages over qubit implementations that require initialization by thermal equilibration and need cryogenic temperatures to operate. The energy difference between ground and excited states must also be large to avoid thermal excitations, avoiding the destruction of spin information.

In addition to the energetic position of the defect states, the occupation of these states is also important. To function as a single-spin center, the defect must be stable in a paramagnetic ground state. As illustrated previously for the NV center (see Fig. 3.6), this involves an exercise in electron counting: placing electrons in states of increasing energy and checking the spin state for each of those configurations. We saw that for the NV center, the desired $S = 1$ state required an overall negative charge state of the center; by fortunate coincidence, the Fermi level in nitrogen-doped diamond (which

is determined by isolated N atoms acting as deep donors) occurs at the right energy to put the NV center in a singly negative charge state. However, for other defects being considered as qubits, one may not be so lucky, and obtaining the correct spin state (and hence charge state) may require separate manipulation of the Fermi level through judicious incorporation of donors or acceptors. An example in the case of SiC is mentioned later in the text (Section 3.3.3).

The electronic structure of the defect must allow for an excited state positioned within the band gap and accessed via a spin-conserving intra-defect optical transition. As alluded to earlier, the presence of a non-spin-conserving decay path can be important for initialization; the nature of these nonradiative transitions is still a subject of active research, which currently renders it difficult to formulate specific criteria that would ensure that such a path be present. Finally, if the qubit is to be probed by the luminescence from an excited state, the transition should be spin-conserving, and the strength of this transition should be large enough to enable efficient, high fidelity measurement of individual defect states.

As for the host material, a wide band gap is desirable if band-to-band and defect-to-band transitions are to be avoided. The host should exhibit small spin-orbit coupling in order to avoid decoherence through spin flips of the defect states. The spin-orbit splitting of the valence-band maximum can be taken as indicative of the strength of the spin-orbit interaction. Also, the constituent elements should have naturally occurring

isotopes of zero nuclear spin, making it possible to eliminate spin bath effects. Host materials that are available in single-crystal form are preferred, either as bulk crystals or epitaxial layers; high structural quality is essential to suppress interactions with point defects or extended defects, or with paramagnetic impurities that could affect the defect spin state. The ability to easily process the host material (e.g., with standard wet etching techniques) is also an asset when considering future device development.

Defects with optical transitions in the near-infrared (0.89—1.65 eV) or visible (1.65—3.10 eV) regions of the spectrum are favored because of the ready availability of optical equipment compatible with these energies. The requirement that defect states lie well away from the band edges then points toward host materials that have sufficiently wide band gaps. Such hosts will typically be fairly ionic and/or contain constituents taken from the first few rows of the periodic table. Diamond, with a band gap of 5.5 eV, clearly fits; and staying with the Group IV elements, SiC is an obvious candidate. Among compounds, oxides tend to have large gaps and may satisfy a number of the other criteria such as zero nuclear spin. Nitrides such as GaN and AlN also have large band gaps, and availability of these materials is rapidly increasing due to their use for solid-state lighting and power electronics. Nitrogen does not have zero nuclear spin, but this is not necessarily a strict requirement.

3.3.2 Paramagnetic defects in SiC: progress to date

Silicon carbide is closely related to diamond in structural and electronic properties. It is tetrahedrally coordinated, every C being surrounded by four Si atoms, and vice versa. A silicon vacancy in SiC is therefore surrounded by four C atoms and may be expected to exhibit behavior very similar to a vacancy in diamond. The most common polytypes, 4H- and 6H-SiC, have a wide band gap of 3.27 eV and 3.02 eV, respectively. The use of defects in SiC for spintronics was proposed by Gali in 2010,[96] and recent experiments have demonstrated that certain defects in SiC exhibit optically addressable spin states with long coherence times.[97–100] Hybrid functional calculations for the Si vacancy and for an NV center in 4H-SiC have indicated that these defects are indeed promising qubit candidates.[92]

Since a Si vacancy in SiC is expected to behave similarly to a vacancy in diamond, placing a N atom next to the Si vacancy should create a center similar to the NV center in diamond. To obtain a spin-triplet ground state, the (N_{C}-V_{Si}) center should be negatively charged. Hybrid functional calculations have indicated that in 4H-SiC, this requires the Fermi level to lie between 1.60 eV and 2.83 eV above the valence band. Unlike the fortuitously favorable situation in diamond, judicious Fermi-level engineering (through controlled doping) is required here to achieve the desired charge state. The calculated configuration coordinate diagram for the NV center in 4H-SiC is shown in

Fig. 3.7). The optical transitions occur at about half the energy of those for the NV center in diamond; for instance, the ZPL of the NV center in 4H-SiC occurs at 1.09 eV, compared with 2.02 eV for NV in diamond. This difference is due to the larger lattice constant of SiC compared to diamond, which leads to a smaller overlap among the carbon sp^3 dangling-bond orbitals, and hence a smaller splitting between the $a_1(2)$ and e levels. This defect has not yet been observed experimentally.

Silicon vacancies have also been detected in various polytypes of SiC. Three zero-phonon lines in the photoluminescence spectrum of 6H-SiC, located in the near infrared between 1.35 eV and 1.45 eV, have been attributed to silicon vacancies at the three inequivalent sites in 6H-SiC: V_{Si} (h), V_{Si} (k1), and V_{Si} (k2); k and h indicate quasi-cubic and quasi-hexagonal sites, respectively. Various first-principles calculations have indicated that V_{Si} incorporates in a high spin state in 3C- and 4H-SiC and may act as a suitable qubit for quantum computing in these and other polytypes of SiC. Specifically, recent calculations show that V_{Si}^{-2} is stable in n-type 4H-SiC and assumes a high spin state, due to the broken tetrahedral symmetry that splits the t_2 states. Calculations also indicate that the charge neutral V_{Si} is a spin-triplet center, stable in p-type and insulating in 4H-SiC (i.e., for Fermi level positions less than 1.4 eV above the valence band). Optical excitation of V_{Si} preferentially pumps the system into specific spin sublevels of the ground state, as experimentally demonstrated by Soltamov *et al.*[101] for 4H- and 6H-SiC.

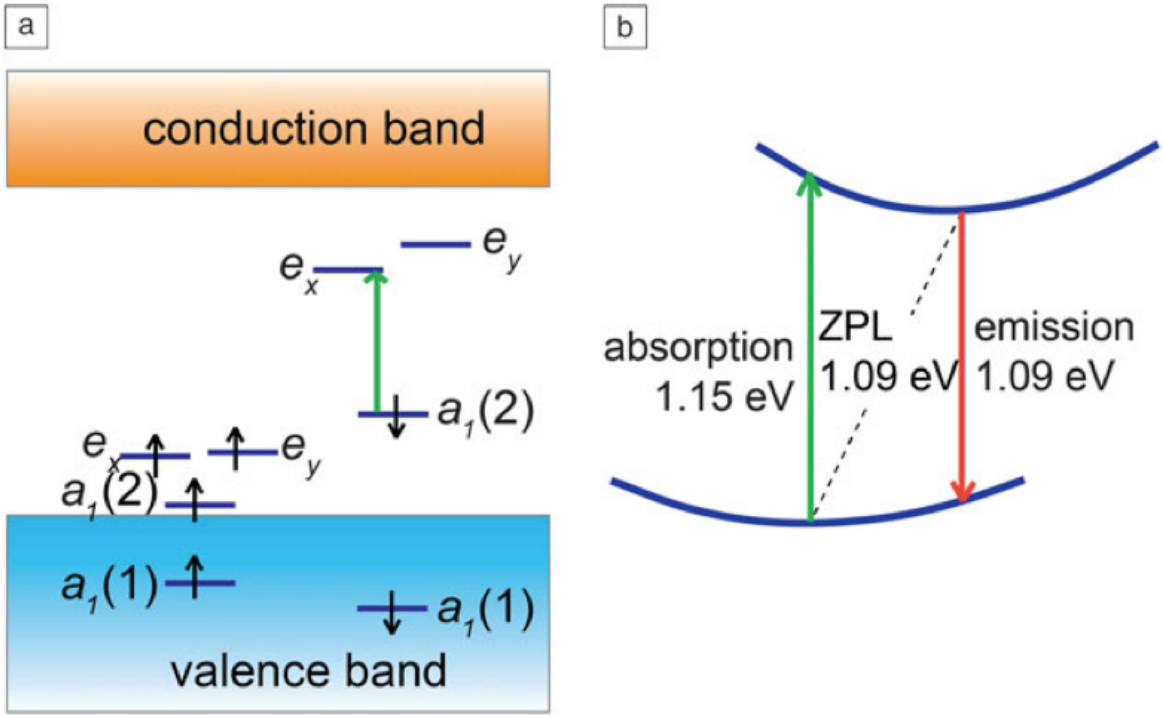

Figure 3.7: (a) Electronic structure of the negatively charged nitrogen-vacancy (NV) center (NV) in 4H-SiC, as calculated with DFT-HSE by my collaborators Justin Weber and Joel Varley in Ref. 92. The positions of the defect states are qualitatively similar to those in the NV center in diamond (Fig. 3.7(a)), but they are located closer to the band edges. Filling the electronic states in order of increasing energy leads to the occupation shown in the figure, resulting in a $S = 1$ (triplet) state for the center. (b) Calculated configuration coordinate diagram for the NV center in 4H-SiC.

The first paramagnetic defect to be coherently manipulated in any polytype of SiC was the divacancy in 4H-SiC.[97] Photoluminescence spectra display a number of distinct zero-phonon lines between 1.09 eV and 1.20 eV. Photo-enhanced spin resonance and annealing experiments have indicated that the first four of these peaks, known by the singular label UD-2, can be attributed to the neutral divacancy. This is a charge-neutral defect complex consisting of a carbon vacancy adjacent to a silicon vacancy. Such a defect can assume four different configurations in 4H-SiC, as illustrated in Fig. 3.8. These have been shown to form localized, paramagnetic electronic states that can be spin-polarized with incident light.

The divacancy signals can be optically detected and coherently controlled. A pulse of light can polarize all the defect spins in the ensemble, and then the ensemble can be coherently manipulated with pulsed microwaves. Excitation with a second pulse of light and measurement of the photoluminescence intensity (a spin-dependent process) allows for a spin-dependent readout of the qubit state. This procedure can be applied to all four inequivalent forms of the divacancy. Very recently, it has been shown that analogous defect centers exist in 3C- and 6H-SiC, which can also be manipulated and optically measured.[98] The microscopic nature of these defects is as yet unknown.

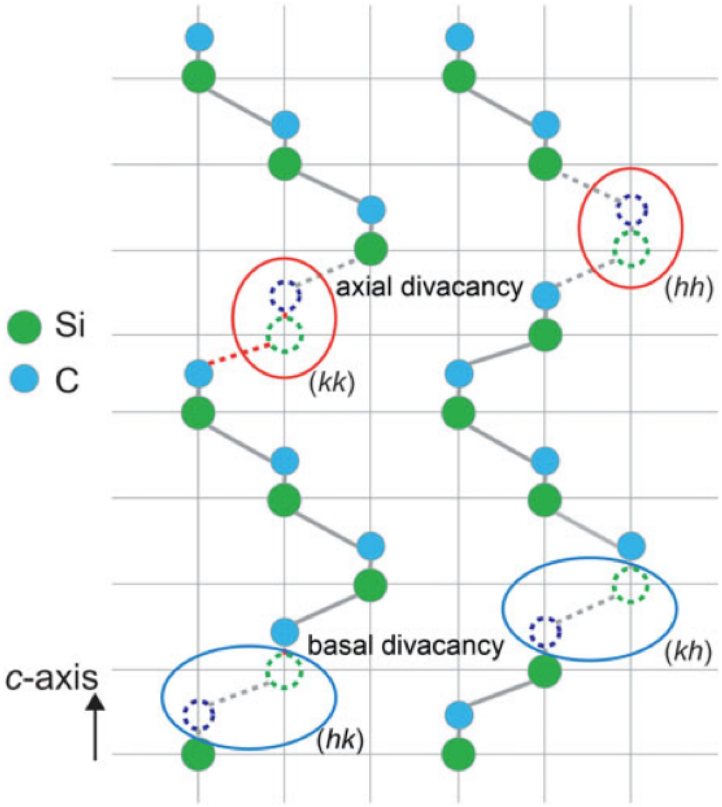

Figure 3.8: Structure of divacancy in 4H-SiC. The complex consisting of a Si vacancy next to a carbon vacancy can occur in four different inequivalent configurations, two axial and two basal. The hh and kk forms of the divacancy are oriented along the c axis of the crystal, while the hk and kh forms are oriented along the basal bond directions.

3.3.3 Finding new qubits in SiC

SiC is a wide-band-gap semiconductor, available as high-quality large single crystals and epitaxial films,[102] and is the basis for a wide variety of electronic and optoelectronic devices.[103] Recent theoretical work has identified the N_C-V_{Si} center in 4H-SiC as a promising candidate for qubit applications.[92] First-principles calculations and experimental measurements have also indicated that the neutral divacancy in 4H-SiC has desirable properties:[96,97,104] photoluminescence, angle-resolved magneto-luminescence, and continuous-wave optically-detected magnetic resonance (ODMR) measurements have demonstrated room-temperature spin control, with spin-coherence properties comparable to those of the NV center in diamond.[97,99,100]

There are four symmetry-inequivalent forms of the divacancy in 4H-SiC, shown in Fig. 3.8. Since all have the same local bonding environment and are similarly stable, they coexist in 4H-SiC, and lead to slightly different excitation and emission energies, which potentially poses problems when trying to address individual centers. Furthermore, the charge-neutral divacancy in 4H-SiC is stable only over a limited range of Fermi-level values within the band gap and requires semi-insulating samples. An alternative to 4H-SiC is to use the cubic polytype 3C-SiC, which has the zinc-blende crystal structure and can also be grown as large single crystals and epitaxial films.[105,106]

The divacancy and NV center are expected to behave similarly in 3C-SiC and 4H-SiC, but the higher symmetry of the 3C polytype eliminates the problem of symmetry-inequivalent configurations. In addition, its smaller band gap (2.36 eV,[107] 0.9 eV smaller than that of 4H-SiC) could potentially be favorable for stabilizing the desired charge states, provided the relevant defect states are not too close to, or resonant with, the valence or conduction bands. A small number of studies have reported on the divacancy in 3C-SiC,[98,108,109] but detailed information about the stability and electronic and optical properties is lacking, providing the motivation for the present study.

Using first-principles calculations we show that the divacancy in 3C-SiC indeed exhibits similar properties as in 4H-SiC, but has the advantage that the desired charge-neutral vacancy with its $S = 1$ spin state is stable over a range of Fermi levels near the conduction-band minimum (CBM) of 3C-SiC, which is easily achieved by n-type doping. In contrast, the neutral divacancy in 4H-SiC is stable only if the Fermi level is between 0.9 and 2.1 eV above the valence-band maximum (VBM), requiring semi-insulating material.

The lattice parameters and band gaps of 3C- and 4H- SiC are obtained using HSE with a mixing parameter of 25% (see Section 1.2.3 for details) along with the experimental values are listed in Table 3.2. We observe excellent agreement between the calculated and experimental values.

Table 3.2: Lattice parameters and band gaps of 4H-SiC and 3C-SiC calculated using the HSE functional. Experimental values[107,110] are listed for comparison.

Material	Property	HSE	Exp.
4H-SiC	a (Å)	3.07	3.07
	c (Å)	10.04	10.05
	E_{g} (eV)	3.17	3.23
3C-SiC	a (Å)	4.35	4.36
	E_{g} (eV)	2.24	2.36

To study the properties of the defects, we perform calculations using a 96-atom supercell for 4H-SiC and a 216-atom supercell for 3C-SiC, with a 2×2×2 Monkhorst-Pack k-point mesh for integrations over the Brillouin zone, and a 400 eV cutoff for the plane-wave basis set. The stability of the different charge states for each defect is determined by calculating formation energies.[7] In the case of a divacancy in 4H-SiC, the formation energy for a given charge state of the defect is given by:

$$E^f[(V_{\mathrm{Si}}\text{-}V_{\mathrm{C}})^q] = E_{\mathrm{tot}}[(V_{\mathrm{Si}}\text{-}V_{\mathrm{C}})^q] - E_{\mathrm{tot}}(\mathrm{SiC}) + \mu_{\mathrm{Si}} + \mu_{\mathrm{C}} + q \cdot \epsilon_F + \Delta^q, \tag{3.3}$$

where $E_{\mathrm{tot}}(\mathrm{SiC})$ is the total energy of the pristine 4H-SiC supercell and $E_{\mathrm{tot}}[(V_{\mathrm{Si}}\text{-}V_{\mathrm{C}})^q]$ is the total energy of the supercell containing the defect in charge state q. μ_{Si} and μ_{C} are the chemical potentials of Si and C referenced to the total energy per atom of bulk Si and diamond, respectively, and can vary over a range limited by the enthalpy of formation of SiC. The Si-rich limit is set by μ_{Si}=E_{tot}(Si) and thus μ_{C}=E_{tot}(C)+ΔH_f(SiC), and

the C-rich limit by μ_{Si}=E_{tot}(Si)+ΔH_f(SiC) and μ_{C}=E_{tot}(C), where ΔH_f(SiC) is the enthalpy of formation of 4H-SiC. In the case of NV centers, we considered Si_3N_4 as the solubility-limiting phase, allowing us to set an upper bound on the N chemical potential. ϵ_F is the Fermi level, which is referenced to the VBM. Δ^q is the correction term to align the electrostatic potential in the perfect bulk and defect supercells and to account for finite-cell size effects on the total energies of charged defects.[8,9]

Constrained DFT can be used to determine excitation energies from first principles calculations. As an example, we can take a $S = 1$ center which has coherent defect levels in the spin minority channel, such as the NV center in diamond, or the divacancy in SiC. In order to calculate the magnitude of the absorption, emission and zero-phonon-line energies between these defect levels, we fix the occupation of Kohn-Sham eigenvalues. An electron in the $a_1(2)$ level can be placed in the e_x level; the atomic structure is fixed, and an electronic relaxation is performed. This allows us to determine an absorption energy. By subsequently allowing the system to relax, we can predict a Frank-Condon shift, and apply this to the absorption energy to determine the zero-phonon line energy (ZPL). The energy of this structure in the ground state can also be determined, allowing us to predict an emission energy. These values can be directly matched with experiment, typically to within 0.1eV.

Using recently developed general techniques from the Van de Walle group,[111] and in collaboration with Audrius Alkauskas, we can precisely calculate luminescence lineshapes

of defects. Starting with a first-principles calculation of the electronic structure of the relevant system, and a subsequent calculation of the phonon modes, we can map the multi-dimensional vibrational problem onto an effective one-dimensional configuration coordinate diagram. This allows the explicit calculation of the normalized luminescence intensity in the Franck-Condon approximation, written as:

$$G(\hbar\omega) = C\omega^3 A(\hbar\omega),$$

where the normalized spectral function A($\hbar\omega$) is defined as:

$$A(\hbar\omega) = \Sigma_n | < \chi_{e0} | \chi_{gn} > |^2 \delta(E_{ZPL} - \hbar\omega_{gn} - \hbar\omega),$$

where C is a normalization factor, χ are the ionic wavefunctions, E_{ZPL} is the zero-phonon line energy (taken from experiment) and the sum runs over all vibrational levels with frequencies ω_{gn}.

3.3.3.1 Stability of charged defects

In Fig. 3.9 we show formation energies for single vacancies, divacancies and NV centers in 4H-SiC and 3C-SiC in the C-rich limit. The slopes indicate the stable charge state for a given position of the Fermi level [see Eq. 3.3], and the kinks represent the charge-state transition levels (q/q'), i.e., the Fermi-level position below which the defect is stable in charge state q and above which it is stable in charge state q'.[7] Note that different charge states for the same defect involve distinct local lattice relaxations.

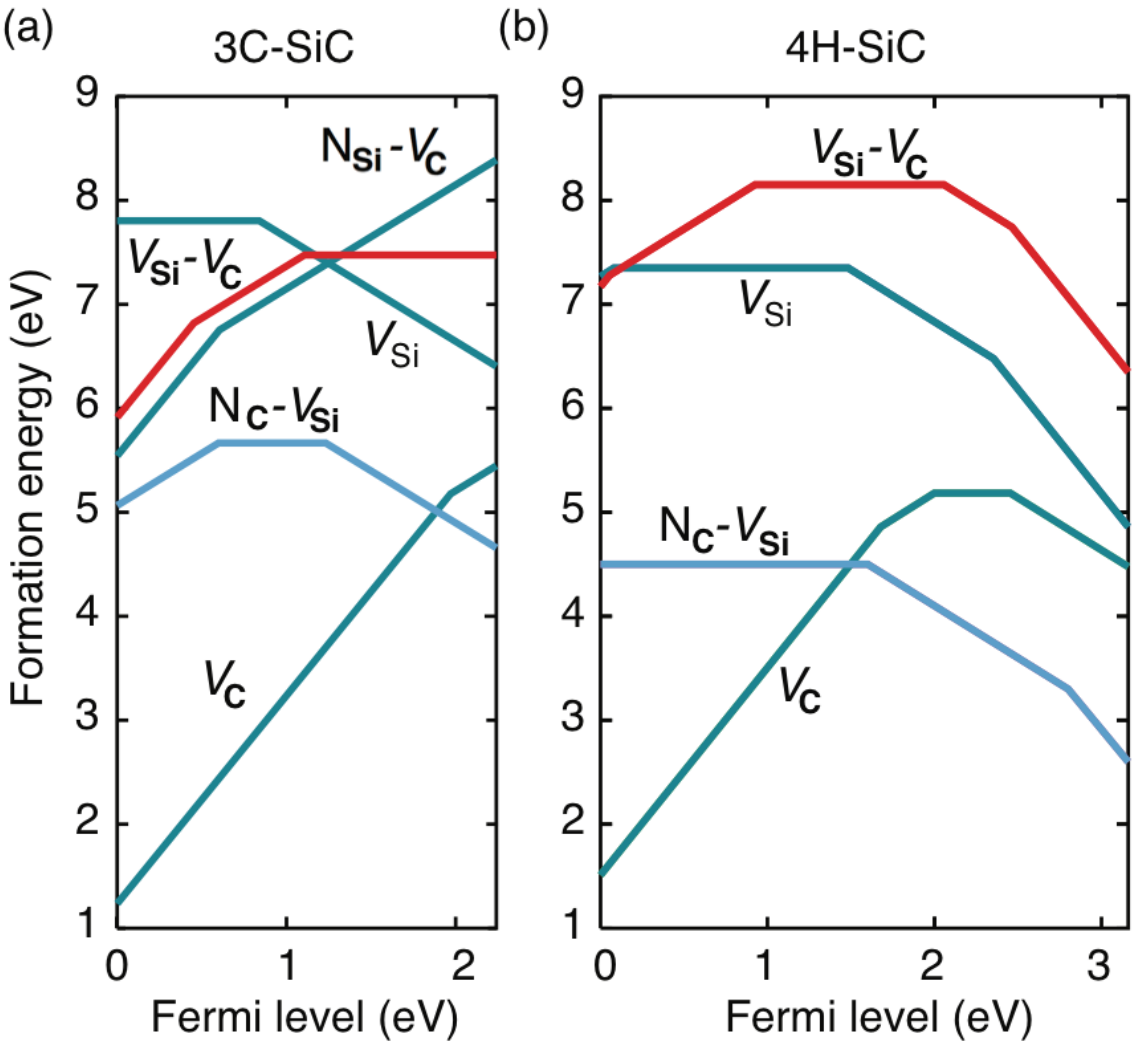

Figure 3.9: Formation energy as a function of Fermi level for single vacancies (V_C and V_{Si}), divacancy (V_C-V_{Si}), and NV centers in 3C- and 4H-SiC, under C-rich conditions.

The formation energies for the divacancies and NV centers are very high and imply that the concentration of these defects, which is proportional to e^{-E_f/k_BT}, are expected to be extremely low under equilibrium conditions. However, as previously observed by Torpo *et al.*,[108] the binding energy of the divacancy is very high, around 3-4 eV for Fermi level positions near the middle of the band gap. Since SiC is grown at temperatures exceeding 1600 °C,[105] single vacancies that are formed during growth will be highly mobile, tending to meet and form complexes. Divacancies can also be intentionally created through irradiation and annealing.[98]

Perhaps more relevant than the absolute formation energies, for the purposes of the present study, are the positions of the transition levels and the range of Fermi-level values over which defects are stable in a given charge state. For instance, the divacancy in 4H-SiC is predicted to be stable in the neutral charge state for Fermi-level values between 0.9 eV and 2.1 eV, requiring 4H-SiC to be semi-insulating. This is consistent with experimental observations.[97] In the case of 3C-SiC, we find that the neutral divacancy is a S=1 triplet and is stable for all Fermi levels above 1.0 eV. In addition, we find that the negative N_C-V_{Si} center also forms a spin triplet and is stable for all Fermi levels above 1.3 eV.

Further insight into the stability of the different charge states of a given defect is provided by inspecting the spectrum and occupancy of the defect-related single-particle states in the gap. Isolated vacancies in SiC introduce levels that are derived from the

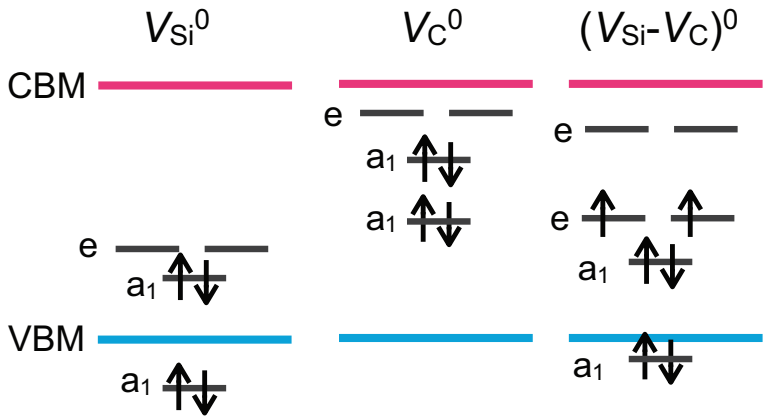

Figure 3.10: Schematic single-particle-based electronic structure of the ground state of neutral Si and C vacancies and the neutral divacancy in 4H-SiC. VBM indicates the valence-band maximum and CBM the conduction-band minimum.

four dangling bonds of the surrounding atoms. In 4H-SiC these dangling bonds combine into two symmetric one-electron states (a_1) and a doubly degenerate e state. In the case of a Si vacancy, the states in the gap are derived from C dangling bonds and lie close to the valence band; for the C vacancy, the states are derived from Si dangling bonds and lie closer to the conduction band, as shown in Fig. 3.10. In the case of divacancies, there are six dangling bonds, which combine into two a_1 and two e states. The lower a_1 and lower e states are derived from the C dangling bonds and the higher a_1 and e states are derived from the Si dangling bonds. In the neutral charge state four electrons occupy the a_1 states, and two electrons occupy the lower degenerate e level. Following Hund's rule, the electrons in the e level have parallel spin.

A key difference between the divacancy in 3C-SiC and 4H-SiC is that the e states associated with the Si dangling bonds lie above the CBM due to the smaller band gap

of 3C-SiC, meaning that it is not possible to occupy them with electrons. As a result, negative charge states cannot be stabilized and the neutral charge state is stable for all Fermi levels in the upper part of the gap. As such, this will be the stable charge state in n-type 3C-SiC. In 4H-SiC, on the other hand, the Si-dangling-bond-associated e states lie in the gap, and the (0/–) and (–/–2) transition levels occur at 1.2 and 0.7 eV below the CBM. Similarly, for the N_C-V_{Si} center the empty spin-down e state, derived from Si dangling bonds, lies above the CBM in 3C-SiC but within the gap in 4H-SiC. As a consequence, the –2 charge state cannot be stabilized for the N_C-V_{Si} center in 3C-SiC, but does occur in 4H-SiC, with a (–/–2) transition level at 0.4 eV below the CBM.

3.3.3.2 Transition energies of paramagnetic defects

In analogy with the NV center in diamond, the centers that are promising as qubits in SiC are those that have an S=1 ground state that can be optically manipulated by exciting an electron from the highest a_1 state to one of the e states, as shown in Fig. 3.5. This occurs for the neutral divacancy $(V_C\text{-}V_{Si})^0$ and the negatively charged $(N_C\text{-}V_{Si})^-$ center. The peak excitation and emission energies can be calculated by constraining the occupation and taking total-energy differences between the excited and ground state configurations, allowing the determination of a configuration coordinate diagram. The calculated configuration diagrams are shown in Fig. 3.11 and the energies of the optical transitions are listed in Table 3.3.

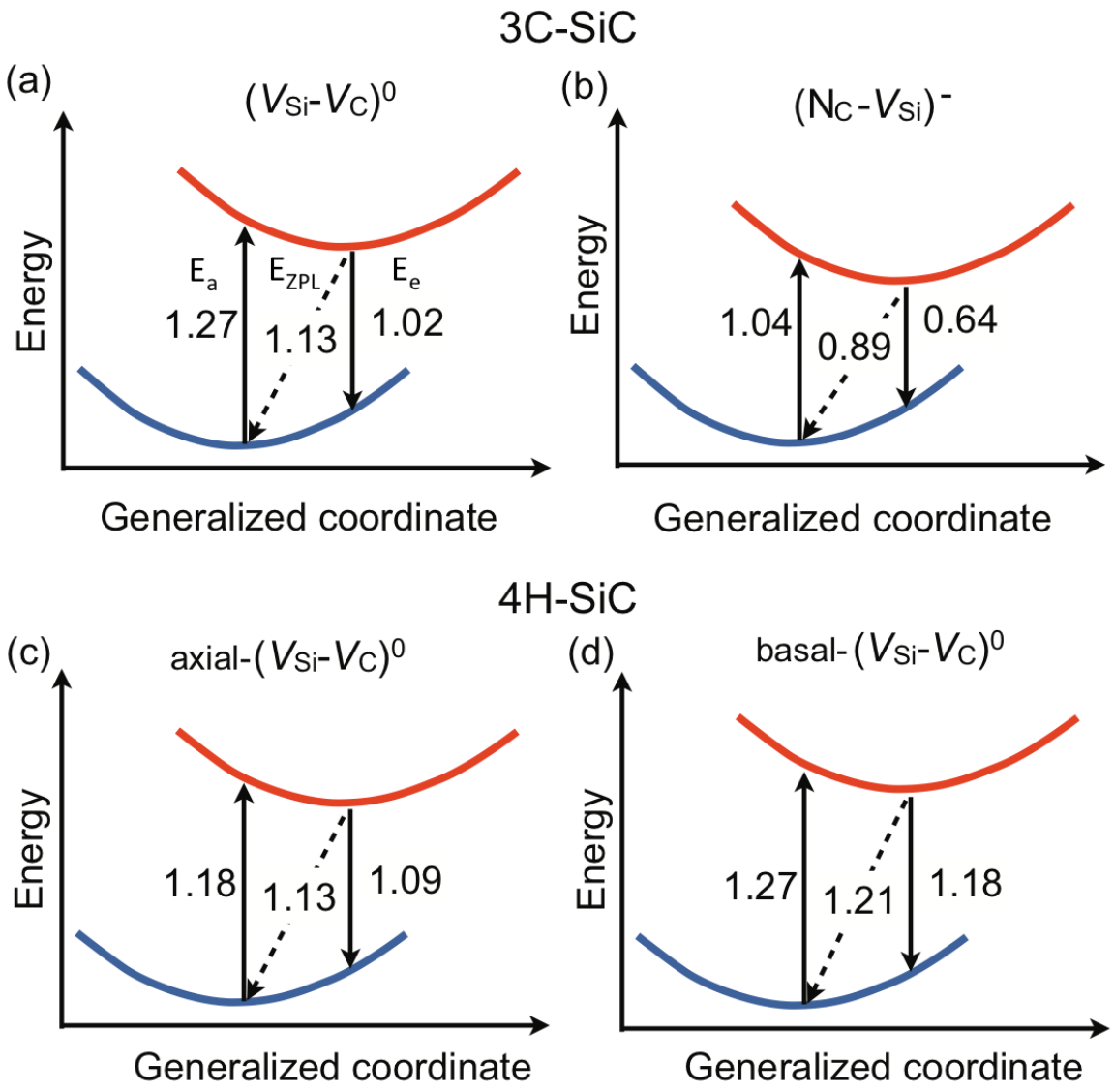

Figure 3.11: Configuration coordinate diagrams for (a) the divacancy and (b) the NV center in 3C-SiC, and for (c) the axial (hh) and (d) the basal (hk) divacancies in 4H-SiC. All values for absorption (E_a), emission (E_e), and zero-phonon line (E_{ZPL}) energies are in eV.

Table 3.3: Calculated absorption, zero-phonon line (ZPL) and emission energies for all divacancies in 4H-SiC, and for the divacancy and N_C-V_{Si} center in 3C-SiC. Experimental values[97,109] for E_{ZPL} are shown where available, although the individual identifications are unique to this paper.

Polytype	Defect	E_a	E_{ZPL}	E_e	E_{ZPL}^{exp}
4H	V_{Si}-V_C (hh)	1.18	1.13	1.09	1.094
4H	V_{Si}-V_C (kk)	1.19	1.14	1.10	1.095
4H	V_{Si}-V_C (hk)	1.27	1.21	1.18	1.120
4H	V_{Si}-V_C (kh)	1.29	1.24	1.21	1.150
3C	V_{Si}-V_C	1.27	1.13	1.02	1.121
3C	N_C-V_{Si}	1.04	0.89	0.64	

In 4H-SiC, there are four different forms of the divacancy: the kk and hh forms of the axial divacancy (k stands for quasi-cubic position and h for quasihexagonal), and the kh and hk forms of the basal divacancy(see Fig. 3.8 for an illustration). We have calculated absorption, emission and zero-phonon line energies for each defect, as shown in Fig. 3.11 and Table 3.3.

The results for ZPLs agree extremely well with experiment,[97] to within 0.1 eV accuracy in all cases. The hh axial divacancy is experimentally observed to be within 0.001 eV of the kk divacancy, consistent with our calculation of a small 0.01 eV difference. We find the hk divacancy to lie 0.03 eV lower in energy than the kh divacancy, in complete agreement with experiment.[97] Based on our calculations, we can assign the hh and kk

axial divacancies to the experimental peaks at 1.094 eV and 1.095 eV respectively, and the hk and kh basal divacancies to the experimental peaks at 1.12 eV and 1.15 eV, respectively.

In the case of the neutral divacancy in 3C-SiC we find an excitation energy of 1.27 eV, an emission peak at 1.02 eV, and the ZPL at 1.13 eV. The divacancy has not yet been conclusively experimentally identified in 3C-SiC. However, an optically detected magnetic resonance study of 3C-SiC by Son *et al.*[109] reported an S=1 defect with trigonal symmetry and a measured ZPL of 1.121 eV. The authors hypothesized that this peak was due to a complex involving a silicon vacancy and another intrinsic defect. We think it highly likely that this defect is in fact the S=1 neutral divacancy. Recent experimental work[98] has confirmed the existence of this spin-1 defect in 3C-SiC with a peak at 1.121 eV, and shown that it is coherent and optically addressable up to room temperature.

The $(N_C\text{-}V_{Si})^-$ center in 4H-SiC has been previously suggested[92] as a potential NV-like qubit. Our calculations show that the $(N_C\text{-}V_{Si})^-$ center in the 3C polytype is also highly promising as a qubit: it is a spin triplet in the ground and excited states, with emission and absorption peaks in the infrared. Infrared luminescence has rarely been used to study NV-like point defects, possibly because the luminescence in that wavelength region tends to be dominated by transition-metal and rare-earth impurities.[112] Vanadium and erbium, in particular, are common impurities in as-grown SiC, and one observes sharply peaked emission spectra between 0.8 and 0.95 eV,[113] arising from intra-

d-shell and intra-*f*-shell transitions. Our calculations yield a binding energy of 2.0 eV for the N_C-$V_{Si})^-$ center in 3C-SiC, consistent with the value obtained in 4H-SiC by Weber *et al.*[114] A nitrogen-doped SiC wafer, annealed at above 700 °C,[104] will lead to the formation of numerous $(N_C$-$V_{Si})^-$ centers, with ZPL energies around 0.89 eV for 3C-SiC, and 1.09 eV[114] for 4H-SiC.

3.3.3.3 Hyperfine coupling parameters

The interaction of the electronic wave functions with nuclear spins can be probed using experimental techniques such as electron paramagnetic resonance (EPR)[115] and related methods such as optically-detected magnetic resonance (ODMR) spectroscopy; the results are expressed in terms of hyperfine parameters. These values provide information about the symmetry of the defect, and the types of atoms involved. However, in order to use this data to make predictions of the microscopic nature of particular defects, highly accurate theoretical comparisons are required. Using a method first described by Van de Walle and Blochl[116] for calculating hyperfine parameters for defects in semiconductors using spin-density functional theory and pseudopotentials, and recently implemented in VASP for use with hybrid functionals, we can calculate hyperfine parameters for arbitrary defects, and make direct comparison with experiment. Accurate calculations of hyperfine parameters require a high-level functional, as calculated param-

eters are very sensitive to the local geometry, and particularly to electron correlations, which is not perfectly treated in traditional DFT functionals.

In Table 3.4, we show calculations for the hyperfine parameters of the neutral hh axial divacancy in SiC, using the PBE and HSE functionals, and comparing with experiment. In general, the agreement is very good: the PBE results seem somewhat closer to experiment than HSE, contrary to expectation; however, this is likely a coincidence. Calculations have also been performed for other defects in SiC and C: specifically V_{C}^{+} and V_{Si}^{+} in SiC, and the NV center. In all cases calculated values match very well with experiment.

3.3.3.4 Photoluminescence lineshapes

Using the technique expounded in Section 3.3.3, we can calculate photoluminescence lineshapes of the four different configurations of the divacancy in SiC, and compare with experiment performed by the Awschalom group at UCSB. Qualitatively, the lineshapes match very well (shown in Figs. 3.12—3.15), as do the relevant lineshape parameters, shown in Table 3.5. The energy difference between the ZPL and the first phonon replica is 30-40 meV in both theory and experiment. The Huang-Rhys factors (related to Debye-Waller factors by the relation $DB = e^{-HR}$) are consistently smaller in theory than in experiment: however, both theory and experiment agree that the Huang-Rhys factors are smaller than observed in the NV center in diamond; this is due to the increased

Table 3.4: Hyperfine parameters of the hh neutral divacancy in SiC.

	A_{xx}	A_{yy}	A_{zz}
Experiment			
C(1-3)	53	50	110
Si(1-3)	3	3	3
Theory (PBE)			
C(1-3)	59-60	58-59	121-127
Si(1-3)	0.7-1.7	0.5-1.5	3
Theory (HSE)			
C(1-3)	73	72	132-146
Si(1-3)	-1.3-0.3	-0.3-0.15	-1.4-1.47

Table 3.5: Lineshape parameters in theory and experiment.

Divacancy	D-W factor	H-R factor	$\hbar\Omega$ (meV)
hh	5%	40	2.99
hh (experiment)	3.4 - 3.6 %	35	3.32 - 3.38
kk	6.7 %	30	2.70
kk (experiment)	3.8 - 4.0 %	34	3.22 - 3.27
hk	5.9 %	30	2.83
hk (experiment)	3.7 - 3.9 %	44	3.24 - 3.30
kh	7.1 %	38	2.64
kh (experiment)	5.5 - 5.9 %	33	2.83 - 2.90

weight of the ZPL relative to the remainder of the line, and could be very significant for increasing the efficiency of photonic structures and heralded entanglement schemes.

3.3.4 Qubits everywhere: possibilities

Over the course of this section, we have surveyed the current status of employing defects as qubits analogous to the NV center in diamond. We have focused primarily on SiC, towards which much of the experimental work is directed. Indeed, numerous defects in 3C and 4H-SiC have been identified, and experiments on optical initialization and control of silicon vacancy and divacancy centers have already demonstrated that

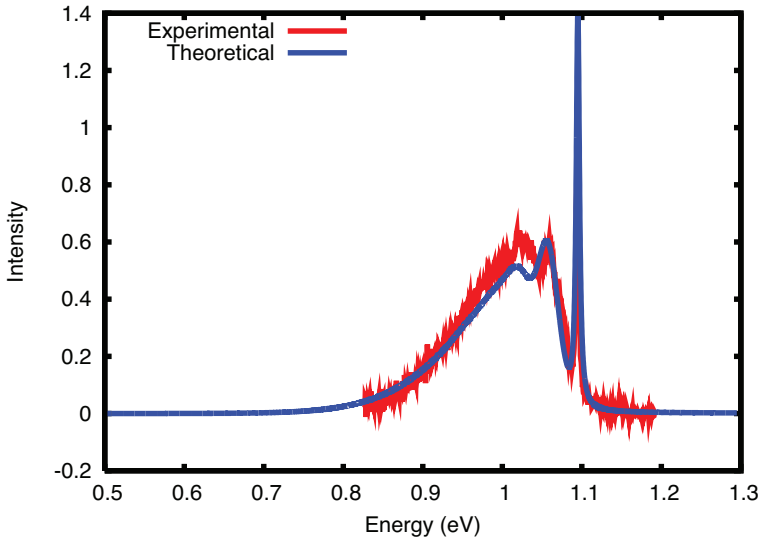

Figure 3.12: Theoretical and experimental photoluminescence lineshape for hh axial divacancy. Experimental measurements provided by A. Falk of the Awschalom group.

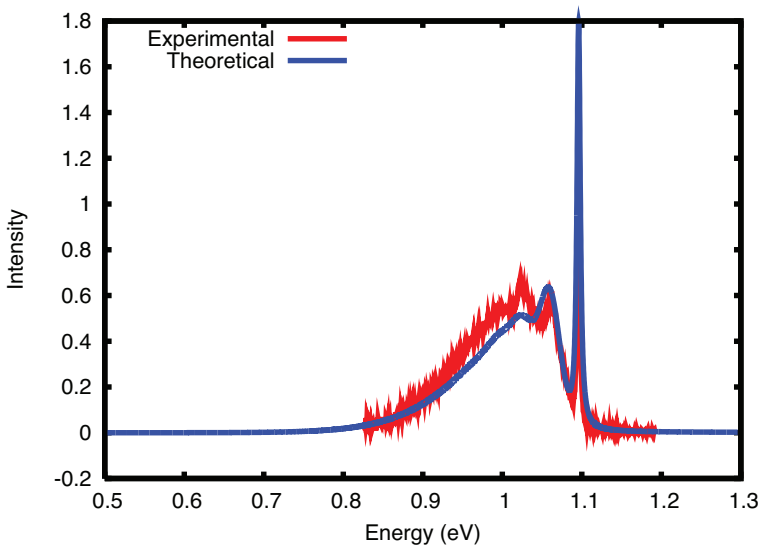

Figure 3.13: Theoretical and experimental photoluminescence lineshape for kk axial divacancy. Experimental measurements provided by A. Falk of the Awschalom group.

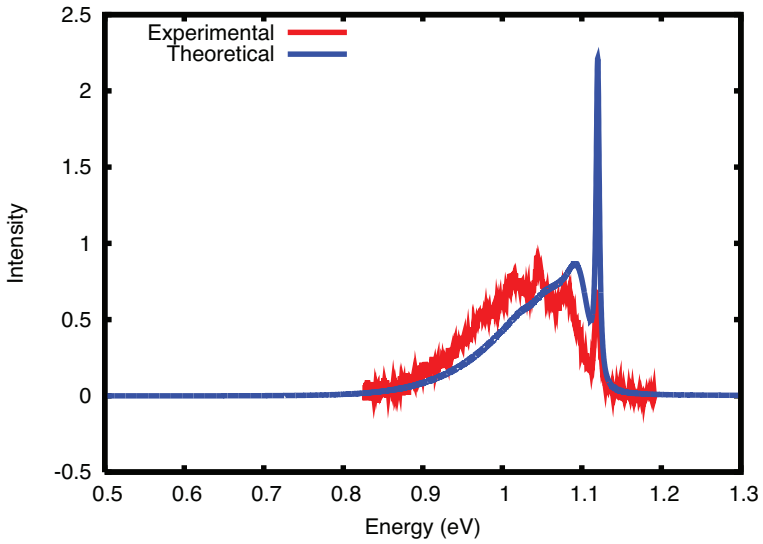

Figure 3.14: Theoretical and experimental photoluminescence lineshape for hk axial divacancy. Experimental measurements provided by A. Falk of the Awschalom group.

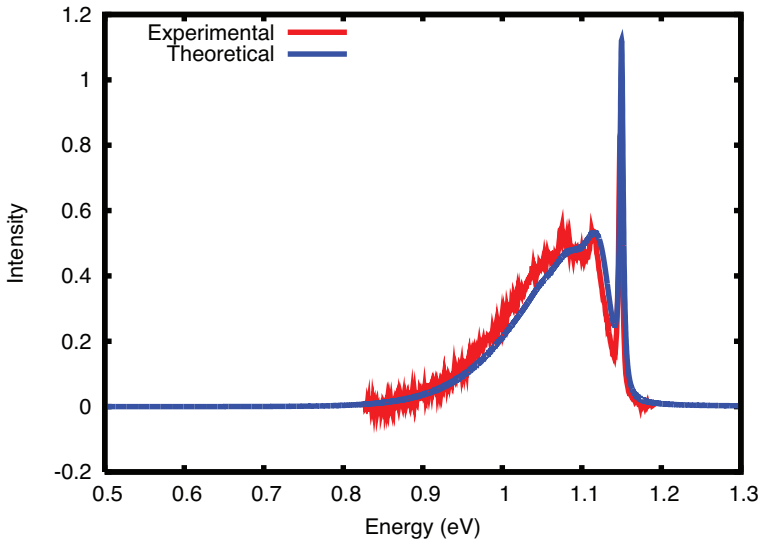

Figure 3.15: Theoretical and experimental photoluminescence lineshape for kh axial divacancy/ Experimental measurements provided by A. Falk of the Awschalom group.

SiC is a promising material. In addition, we have used first-principles calculations to study properties of various point defects in 4H-SiC and 3C-SiC that can serve as qubits, reporting their formation energies, transition levels, and optical transition energies. Our calculations for the divacancy in 4H-SiC allow us to assign the experimentally observed ZPLs[97] to specific microscopic vacancy configurations. Our results for the neutral divacancy in 3C-SiC suggest that this is the defect that has been observed to give rise to a spin-1 center with a ZPL peak at 1.121 eV.[98,109] We also find that the $(N_C\text{-}V_{Si})^-$ center in 3C-SiC is highly promising for qubit applications. Unlike the situation in 4H-SiC, where semi-insulating material is needed with the Fermi level within a specific range, the desired S=1 state of the $(V_C\text{-}V_{Si})^0$ and the $(N_C\text{-}V_{Si})^-$ centers is stable in n-type-doped 3C-SiC, which should make experimental observation and control more straightforward than in the 4H polytype. In addition, we have calculated hyperfine coupling parameters and photoluminescence lineshapes for the neutral divacancy in 4H-SiC, in excellent agreement with experiment.

Other than those in SiC, several new defect-based qubits have recently also been computationally investigated: Chanier *et al.*[117] have identified substitutional nickel impurities in diamond as potential qubits, and Yan *et al.*[118] have proposed the neutral V_{Ga}—O_N complex in zinc-blende GaN. A vast range of defect centers in other materials remains to be explored. In order to effectively create and manipulate these qubits, better control over single defect centers must be achieved. Optical polarization has been re-

ported for vacancy-related complexes in the cubic semiconductors MgO and CaO,[119,120] which may herald a new path to defect-based qubits in octahedrally coordinated semiconductors. In addition, while the emphasis in the present article was on vacancy-related centers, other point defects (such as antisites) as well as impurities can be investigated in the same fashion. Chromium-implanted diamond leads to a fully-polarized single-photon emitter with a sharp ZPL at 1.66 eV;[121] however, the local structure of this defect is currently unknown. Substitutional transition metals may be a promising avenue for research since they tend to incorporate in high-spin states; in fact, a six-level system based on Mn has been realized in MgO.[122] Investigations on these and other centers offer great potential for qubit functionalization, but will also deepen our understanding and insight in this fascinating area of physics.

3.4 Conclusions

In reaching the goal of a functioning quantum computer, understanding the role of atomic-scale defects is essential. First, the contribution of defects to decoherence in superconducting qubits based on Josephson junctions has been studied, and hydrogen identified as a major source of TLS decoherence in these qubits. The specific tunneling frequencies and qubit-TLS coupling parameters have been calculated and compared with experiment to a high degree of accuracy, and possible ways of mitigating the effect of

TLS have been suggested. Second, specific defects have been proposed as qubits in new materials, such as SiC, analogous to the NV center in diamond. The formation energies, excitation and emission energies, photoluminescence lineshapes, and hyperfine spectra of these defects have been calculated, and compared to experiment where available. This allows us to accurately predict the properties of a range of defects in SiC, and to outline a path towards identifying new defects as qubits in different materials.

Chapter 4

Beyond a cubic nanometer: macroscopic simulations of devices

Deutsch's Law: Every problem that is interesting is also soluble.

David Deutsch

4.1 Introduction

Over the past two chapters, we have demonstrated the power of a first-principles approach in addressing device problems, and in enabling advanced future devices. However, despite the accuracy of density functional theory (DFT), we are still subject to some

fundamental limitations. Specifically, simulating systems greater in size than a few hundred atoms is impossible within the current paradigm. A few hundred atoms is enough to determine optical properties, formation energies, and a host of other properties, but if we want to test our predictions on actual devices, we must extend our calculations to a larger, macroscopic length scale. In this case, we can perform Schrödinger-Poisson calculations, informed by first-principles and, in some cases, experimental input, to accurately model the behavior of devices at scales reaching the 100 nm level, which is completely intractable using DFT.

In this chapter, we will provide examples of recent calculations we have performed for devices using macroscopic device simulations, largely in conjunction with first-principles calculations, and will illustrate the power of using a multi-scale approach to accurately model oxide and nitride-based heterostructures. First, we will present work on the origin of the two-dimensional electron gas (2DEG) at the $SrTiO_3/LaAlO_3$ (STO/LAO) interface, the impact of a field-dependent dielectric function on said 2DEG, and the possibility of using alternative oxide heterostructures to improve the performance of future oxide-oxide devices (Section 4.2). Next, we will investigate the impact of surface donor states at the AlGaN/GaN interface, and study the macroscopic effects caused by atomic-scale hole traps in AlGaN/GaN high-electron-mobility transistors (HEMTs) (Section 4.3).

Results on the origin of the 2DEG in STO/LAO heterostructures were published in Physical Review B,[123] and manuscripts on all other results for oxide heterostructures are in preparation.[124,125] All results on the impact of surface donor states at the AlGaN/GaN interface were performed in collaboration with Maosheng Miao, and were published in J. Phys. D,[126] and a manuscript on hole traps in AlGaN/GaN HEMTs is in preparation.[127] In this chapter, only the Schrödinger-Poisson simulations were performed by the author.

4.2 Oxide-based devices

After the groundbreaking discovery of a high-density two-dimensional electron gas (2DEG) at the interface between two insulators, $SrTiO_3$ and $LaAlO_3$,[128] an investigative push began to explain and repeat this result. The high-density electron gas has been explained as a physical outcome of the polar discontinuity present at the interface between a polar and non-polar material, which provides 1/2 of an electron per unit cell.[123] This implies a maximum sheet carrier density of $3.3\text{x}10^{14}$ cm^{-2}, which has indeed been observed at the $SrTiO_3/GdTiO_3$ interface.[129] Several other oxide interface combinations have been experimentally investigated,[130,131] and many more have been predicted.[132] Such an interface could be used in novel oxide-only devices analogous to heterostructure field effect transistors,[133–135] but with an order of magnitude higher charge density than the current state-of-the-art. In addition, the Mott transitions present in many of these

interfaces may lead to the development of Mott-based field-effect transistors[46] with large on/off ratios and femtosecond switching speeds.[136]

The 2DEG at the STO/LAO interface exhibits densities that are difficult to achieve in conventional semiconductors[128,137,138] and displays unique behavior including ferromagnetism,[139] superconductivity,[140] and even the puzzling coexistence of both.[141] It has been proposed as the basis for novel electronic devices that exploit strong electron-electron correlation in the narrow bands derived from d states of the transition metal.[142] STO/LAO heterostructures have been fabricated using methods that allow unprecedented control over layer thickness, such as pulsed layer deposition (PLD) and molecular beam epitaxy (MBE).[137,138]

4.2.1 A two-dimensional electron gas at the $SrTiO_3/LaAlO_3$ interface

While great progress has been made in characterization and exploitation of the physical phenomena, the mechanisms that determine the density of electrons in the 2DEG have remained a subject of intense debate.[143–149] This lack of understanding inhibits achieving the control that is required for device applications.

Typically, an LAO layer of less than 20 nm is deposited on a TiO_2-terminated [001]-oriented STO substrate or epilayer,[137,138] as shown in Fig. 4.1. The electrical conduc-

tivity at the buried interface is then probed as a function of temperature.[137,138] Carrier densities up to $2\times\ 10^{13}$ cm^{-2} have been reported.[144,150–152] Low-temperature sheet resistance varying from 10^{-2} to 10^{4} has been observed, displaying a strong dependence on the oxygen partial pressure in the growth environment or post-growth annealing treatments.[139,153] A dependence of the sheet resistivity and carrier density on the thickness of the LAO layer has also been observed: heterostructures with LAO layers less than 4 unit cells ($\sim$1.6 nm) thick exhibit insulating behavior, while thicker layers become conducting.[152] However, experiments by Huijben *et al.*[154] indicated the existence of conducting interfaces for LAO layers as thin as two unit cells in the presence of an STO layer on top of the LAO.

4.2.1.1 Where do the electrons come from?

A variety of models have been put forth to explain the origin of the carriers at the interface.[143–149] However, none can account for all the experimental observations, raising questions about their validity and general applicability. The prevailing models assume that the carriers originate from somewhere other than the ideal interface, such as from the supposedly negatively charged LAO top surface (driven by the "polar catastrophe"),[145,146] from oxygen vacancies,[144,147–149,151,153] or from Sr-La intermixing at the interface.[155–157]

Here we note that the source of the carriers should not be in question. As explained in the joint work of Ref. 123 and illustrated in Fig. 4.1, the polar discontinuity at the STO/LAO interface provides an intrinsic source of electrons with a density of 0.5 electrons (e^-) per unit cell (3.3×10^{14} cm^{-2}). In an ionic picture (based on Sr^{2+}, La^{3+}, Ti^{4+}, and O^{2-}), $LaAlO_3$ can be regarded as a stacking of $(LaO)^+$ and $(AlO_2)^-$ planes along the [001] direction; LaO planes donate electrons which become bound in the neighboring AlO_2 planes. $SrTiO_3$, on the other hand, is composed of alternating charge-neutral $(SrO)^0$ and $(TiO_2)^0$ planes. At the interface with TiO_2-terminated STO, the TiO_2 layer is already charge-neutral, causing the interfacial LaO layer to act as a sheet of donors, donating $0.5e^-$ per unit cell, which due to the large conduction-band offset between LAO and STO should flow into the STO. The question is therefore not "where do the carriers come from?", but rather, "where do the electrons disappear to?" I.e., why is the observed carrier density an order of magnitude lower[139,144,150,151] than the expected $0.5e^-$ per unit cell (3.3×10^{14} cm^{-2})?

Based on advanced first-principles calculations performed by my collaborators Lars Bjaalie and Anderson Janotti, combined with Schrödinger-Poisson (SP) simulations performed by the author, we[123] attribute the problem to the lack of a suitable termination for the top surface of the LAO layer. The asymmetric nature of the resulting layer structure drains electrons away from the 2DEG, and further exposes the heterostructure to the detrimental effects of point-defect formation. Our insights enable us to propose

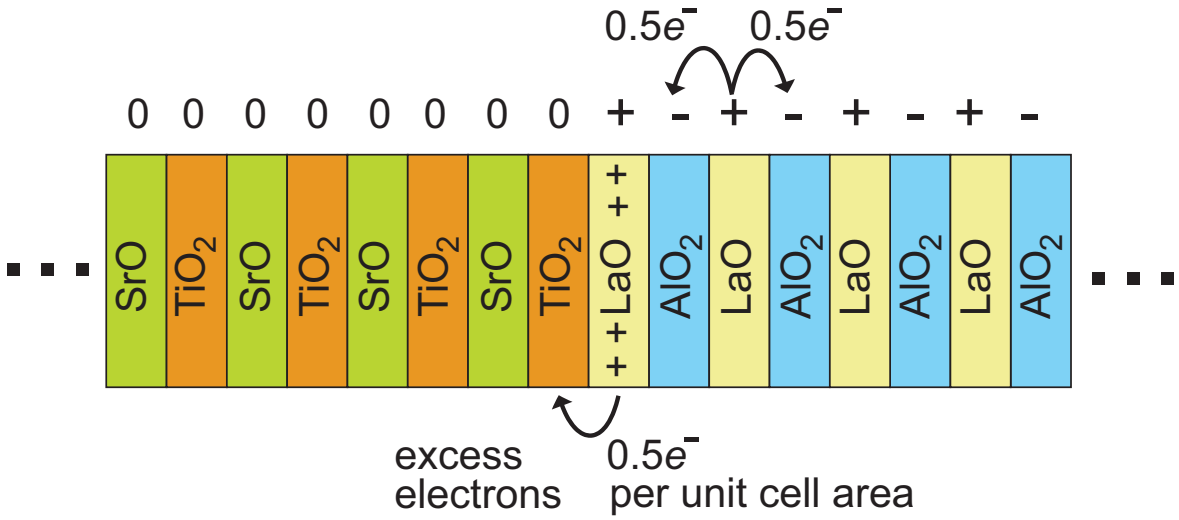

Figure 4.1: Layer structure of a $SrTiO_3/LaAlO_3$ heterostructure with TiO_2-LaO planes at the interface. Nominal charges are indicated above each layer. LaO planes act as electron donors; the TiO_2 plane terminating STO is already charge neutral, and therefore the interfacial LaO plane acts as a delta-doped layer of donors with a density of 0.5 electrons per unit cell. From Ref. 123.

specific strategies for overcoming the problems associated with the current STO/LAO structures, and also provide guidelines for the choice of other materials combinations. We note that 2DEG densities as high as 3.3×10^{14} cm^{-2} have already been observed at $SrTiO_3/GdTiO_3$ interfaces,[158] consistent with the predictions of our model.

The heterostructures were first modeled by my collaborators Lars Bjaalie and Anderson Janotti, using an $(STO)_8/(LAO)_8$ superlattice containing two equivalent TiO_2-LaO interfaces [Fig. 4.2]. The in-plane lattice constant was fixed to that of STO, representing a heterostructure coherently grown on an STO substrate. Full relaxation was allowed for both the out-of-plane lattice constant and all atomic positions.

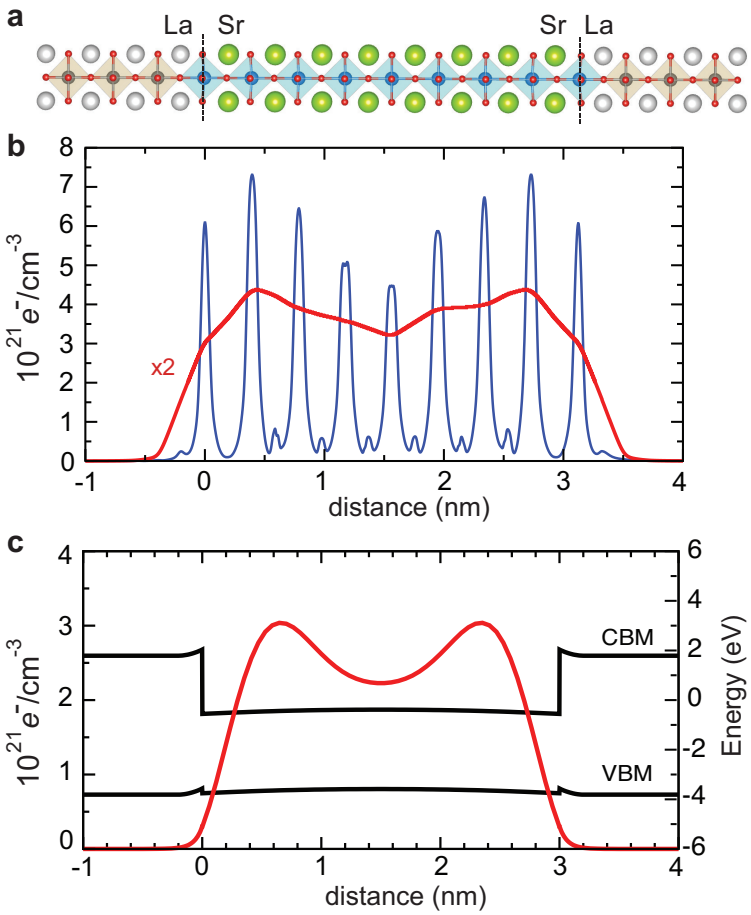

Figure 4.2: First-principles and Schrödinger-Poisson results for a $(SrTiO_3)_8/(LaAlO_3)_8$ superlattice with TiO_2-LaO interfaces. **a**, Atomic structure, with oxygen atoms shown in red. Ti-centered octahedra are shown. **b**, First-principles planar and macroscopically averaged charged density of the occupied subbands. **c**, Schrödinger-Poisson simulations for the same superlattice, showing good agreement with the first-principles results. CBM stands for conduction-band minimum and VBM for valence-band maximum. First-principles calculations were performed by my collaborators, Schrödinger-Poisson calculations were performed by the author, and all results are published in Ref. 123.

The first-principles calculations quantitatively confirm that the free carriers in the 2DEG at the interface originate from the interfacial LaO plane. Fig. 4.2 shows the

integrated charge density, plotted along the [001] direction, associated with the electrons generated at the interface. This charge density appears exclusively on the STO side and corresponds to occupied subbands in the conduction band. The integrated charge is 3.3×10^{-14} cm^{-2} per interface (i.e., $0.5e^-$ per unit cell area), exactly what we expect based on the consideration of the interfacial LaO as a delta-doped donor layer. We observe that this charge is delocalized over multiple planes of Ti and *not* localized on a single Ti layer at the interface. If the latter were the case,[146] the electrons would be immobilized on interfacial Ti^{3+}, which would be inconsistent with the observation of a high-mobility 2DEG.

The symmetric nature of the charge density in the STO layer arises from the fact that the first-principles calculation needs to maintain periodicity along the [001] direction, and therefore corresponds to a superlattice containing two interfaces. It has been verified,[123] however, that the results described here do not depend on the thickness chosen for the individual STO and LAO layers.

Very similar results are produced by SP simulations[11] for the $(STO)_8/LAO_8$ superlattice. The nominal charge density of $0.5e^-$ per unit cell area, as obtained from first-principles calculations,[123] is reproduced in the SP simulation for the $(STO)_8/LAO_8$ superlattice, as shown in Fig. 4.2(c). While these simulations do not include the intricacies of the STO conduction-band structure, they do accurately capture the overall carrier density and distribution near the interface, which is the focus of our study. The

SP simulations allow modeling systems with larger dimensions and, more importantly, lacking periodicity, and hence enable us to study layer structures that are beyond the capabilities of the first-principles calculations.

4.2.1.2 Where do the electrons go?

Given that our simulations so far indicate that the 2DEG density should be expected to correspond to $0.5e^-$ per unit cell, we now address the question of why experiment shows much lower densities.[139,144,150,151] The answer lies in the fact that the type of symmetric structure depicted in Fig. 4.2, with two identical interfaces, is never achieved experimentally. In most practical implementations, the LAO layer is of finite thickness and has a surface terminated on an AlO_2 plane. The consequences are examined in Fig. 4.3 for a 2-nm thick LAO layer (about 5 unit cells) on thick STO. In panel a it is assumed that a perfect STO layer (with TiO_2-LaO planes at the interface) can be deposited on top of the LAO. This effectively reproduces the symmetric situation that was investigated in Fig. 4.2, the only difference being the smaller LAO thickness and larger STO thickness; reassuringly, the results are very similar, with the full density corresponding to $0.5e^-$ per unit cell appearing in the 2DEG. Note that the electrostatic potential (reflected in the slope of the conduction band) is essentially flat across the LAO layer, indicating no charge is being transferred between the two interfaces.

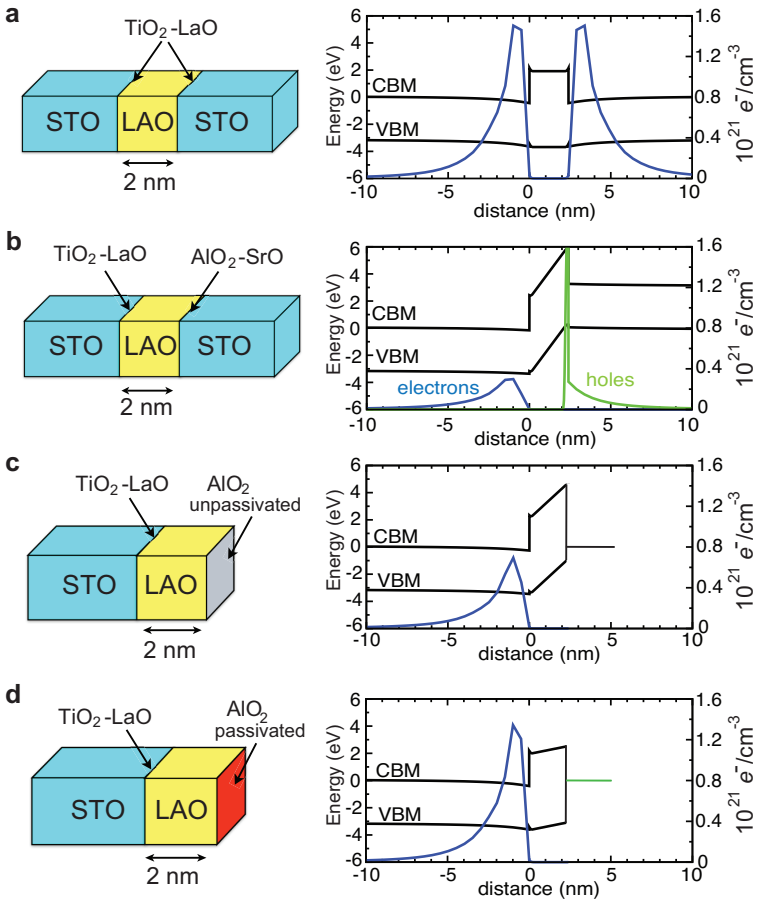

Figure 4.3: Schrödinger-Poisson simulations for $SrTiO_3/LaAlO_3$ interfaces. Layer structures are depicted on the left, and the corresponding band diagrams and charge density distributions on the right. The zero of energy is placed at the Fermi level. **a**, STO/LAO/STO with two equivalent TiO_2-LaO interfaces. The integrated electron density is 3.3×10^{14} cm^{-2} per interface. **b**, STO/LAO/STO with inequivalent interfaces: TiO_2-LaO on the left, AlO_2-SrO on the right. **c**, STO/LAO with TiO_2-LaO at the interface and an AlO_2-terminated surface containing acceptor-like surface states (green horizontal bar). **d**, STO/LAO with TiO_2-LaO at the interface and a passivated surface.

The case depicted in Fig. 4.3(a) is unrealistic, because experimentally it has turned out to be difficult (or even impossible) to grow STO/LAO/STO structures with an interface between LaO and TiO_2 planes on the right-hand side.[154] Fig. 4.3(b) depicts the situation for an STO/LAO/STO layer structure with an interface between AlO_2 and SrO on the right. By similar logic as applied to the TiO_2-LaO interface leading to donor doping, an AlO_2-SrO interface leads to acceptor doping with a sheet density of 3.3×10^{-14} cm^{-2} since the AlO_2 layer is lacking $0.5e^-$ per unit cell which (in the bulk) would come from an LaO plane (the SrO plane in STO being charge-neutral). This asymmetry leads to a strong dipole being set up across the LAO layer (note the slope in electrostatic potential), with electrons transferred from the TiO_2-LaO interface to the AlO_2-SrO interface. In the process, the 2DEG density is drastically reduced, although not to zero because the electric field in the LAO layer results in the valence band of LAO rising above the Fermi level (with a hole gas appearing at the right-hand interface), thus limiting the transfer of electrons.

Having an ideal AlO_2-SrO interface on the right (Fig. 4.3(b)) is similar to terminating the LAO layer with an ideal AlO_2 surface—indeed, all planes in the STO on top are charge-neutral and hence do not contribute to any charge exchange. The only difference is that a realistic AlO_2-terminated LAO surface would exhibit oxygen dangling bonds, giving rise to partially filled surface states with energies in the lower part of the band gap. Similar charge transfer would occur as described for the case depicted in Fig. 4.3(b),

except that the acceptor states are now *deep* acceptor states (with an ionization energy of 1 eV above the LAO VBM), as shown in Fig. 4.3(c). The Fermi level at the surface will be pinned at these acceptor states, limiting the rise of the LAO VBM. This is actually beneficial since it limits the amount of electron transfer out of the 2DEG, compared to the situation of Fig. 4.3(b). Note that this corresponds to a decrease in the slope of the potential, i.e., the magnitude of the electric field, across the LAO layer: the smaller the field, the higher the 2DEG density.

4.2.1.3 How do we keep the electrons?

This suggests a strategy for increasing the electron density in the 2DEG, namely minimizing the slope in the potential across the LAO. Fig. 4.3(d) depicts a fully passivated surface, i.e., the density of acceptors at the surface is assumed to be zero. While there is still a slight slope in the potential (due to the fact that the centers of gravity of the positive and negative charge distributions do not coincide), the 2DEG density now recovers its nominal value of $0.5e^-$ per unit cell.

Our findings easily explain the experimentally observed dependence of 2DEG density on LAO thickness.[152] As noted in the discussion of Fig. 4.3(b) above, full transfer of electrons to the top surface will tend to occur if the LAO layer thickness is insufficient to bring the VBM at the surface above the Fermi level. Given a positive sheet charge density of 3.3×10^{-14} cm^{-2} at the STO/LAO interface, Gauss' law predicts a field of

about 0.25 V/Å, and thus it takes a "critical thickness" of about 3 or 4 unit cells of LAO to develop enough of an increase in potential to bring the VBM (or acceptor-like surface states) of LAO above the Fermi level, at which point transfer of electrons out of the 2DEG is suppressed and observable mobile charge appears in the 2DEG. This is confirmed by explicit SP simulations as a function of LAO layer thickness (not shown). All of this is completely consistent with results in the literature; our main point is that complicated arguments (e.g., relating to interfacial reconstructions) are unnecessary to explain the experimental observations.[152] If the exposed AlO_2 surface is passivated, our model predicts that electron transfer is suppressed and a 2DEG can in principle be observed, even for LAO layers below the critical thickness. This explains why some experiments have observed conducting interfaces for LAO thickness of less than 4 unit cells.[154]

We now turn to the observed variation in 2DEG density with oxygen partial pressure during growth or annealing.[150,153] LAO layers are typically AlO_2-terminated and exhibit partially filled oxygen dangling bonds. Given the high electronegativity of oxygen and the position of such dangling-bond states close to the VBM of LAO, a strong driving force exists to fill these surface states with electrons. Modifications of the surface that remove oxygen dangling bonds suppress electron transfer to the surface, and hence lead to a higher 2DEG density. Growth or annealing in an environment with low O_2 partial pressure results in surface reconstructions containing oxygen vacancies

(or equivalently cation adatoms), effectively removing oxygen dangling-bond states from the surface. The situation is then closer to the scenario of the passivated surface described in Fig. 4.3(d), which shows that a high-density 2DEG can develop. While such oxygen treatments cannot be expected to lead to full passivation, the predicted trend of lower oxygen pressure resulting in higher 2DEG density is definitely consistent with experiment, without having to invoke modification of or point-defect formation at the buried STO/LAO interface. Indeed, we consider the latter unlikely due to the high formation energy of oxygen vacancies in n-type STO.[148]

Conversely, annealing under high O_2 partial pressure leads to a higher density of oxygen-related surface states that will consume electrons from the 2DEG. This explains the seemingly puzzling fact that attempts to perfect the structural quality of STO/LAO heterostructure by growing or annealing under high O_2 partial pressure often leads to high sheet resistance or insulating behavior at the interface.[137]

Manipulating oxygen partial pressure may not be the most effective means of passivating the LAO surface. Hydrogen tends to be a good passivating agent. Indeed, first-principles calculations[159] found that hydrogenation of the LAO surface leads to an increase in the 2DEG density, although we disagree with Son *et al.*'s interpretation[159] that hydrogen donates electrons to the interface. The correct picture, as argued above, is that hydrogen passivates the surface and in the process prevents electrons being drained

away from the interface. Experimental efforts to identify the most effective means of controlling and passivating LAO surfaces could be highly fruitful.

Another approach to prevent electron transfer to the surface is to provide a source of electrons to the surface, for instance by depositing a metal on top of the LAO. The effect will depend on the metal used, specifically, on the work function of the metal relative to the electron affinity of STO. Metals with work function larger than the electron affinity of STO (e.g., Au) will still result in suppressed 2DEG densities since they will not succeed in suppressing electrons draining away from the interface. Metals with work functions equal to or smaller than the electron affinity of STO (such as Ti or Al) are needed to increase 2DEG densities. These qualitative insights are confirmed by explicit SP simulations. These effects have also been observed in recent first-principles calculations,[160] although again the interpretation should not be that electrons are being transferred from the metal to the interface. Ultimately, though, metal capping layers on STO/LAO heterostructures may be more of academic interest, since it may prevent experimental probing of the 2DEG and also prove incompatible with device applications. Emphasis on surface passivation techniques, as described above, is a more promising route.

Finally, we note that our model of the fundamental physics at STO/LAO interfaces, developed in collaboration with Lars Bjaalie and Anderson Janotti,[123] transcends the specific materials system being discussed here and is generally applicable to oxide inter-

faces. For instance, it explains why a 2DEG with the full nominal density of 3.3×10^{-14} cm^{-2} has been observed at $SrTiO_3/GdTiO_3$ interfaces.[158] $GdTiO_3$ (GTO) is composed of alternating positively charged $(GdO)^+$ planes and negatively charged TiO_2^- planes, similar to $LaAlO_3$ (note that Ti has valence 3 in GTO). An STO/GTO interface with TiO_2-GdO interfacial planes will therefore also act as a sheet of donors. The difference with the STO/LAO case lies in the fact that no electric field occurs within the GTO layers. Indeed, STO can be grown with high quality on top of GTO, and since the interfaces are always between TiO_2 and GdO planes the top and bottom interfaces of each GTO layer are identical by construction. This symmetry prevents an electrostatic potential buildup, as shown in Figs. 4.2 and 4.3(a) and allows the full 3.3×10^{-14} cm^{-2} density to be present in the 2DEG.

Even in the absence of an STO overlayer, a buildup of electrostatic potential is unlikely in the GTO layer. If GdO-terminated, both the interface and surface would exhibit donor-like behavior. If TiO_2-terminated, the electrons that are needed to fill acceptor states do not flow into deep-lying oxygen-derived states, but rather into a Ti-derived lower Hubbard band which lies not far below the CBM of GTO.[158] This position of the Fermi level at the surface again suppresses a buildup of potential and maintains the full 2DEG density at the STO/GTO interface.

In summary, based on first-principles calculations and Schrödinger-Poisson simulations, and in collaboration with my coauthors,[123] we have demonstrated that electronic

conductivity at the STO/LAO interface arises from electrons that are intrinsic to the interface. This precludes the need to invoke other sources of electrons such as the top LAO surface (according to the polar catastrophe model[145,155]), or oxygen vacancies acting as donors.[144] The suppression of the 2DEG density at STO/LAO interfaces has often been attributed to interfacial reconstructions (either atomic or purely electronic, based on mixed valence of Ti), which in turn were invoked as a consequence of a "polar catastrophe". Our present results show there is no need for invoking such mechanisms. They emphasize the need for measures to prevent electrons draining away form the interface, which can be accomplished by preventing an electrostatic potential buildup in the LAO layer. Proposed strategies include passivation of the surface, or depositing metals with suitably low work functions. These insights into the origin of carriers at the STO/LAO interface will pave the way to enhanced control of the 2DEG at the interface of complex oxides.

4.2.2 A variable dielectric function in $SrTiO_3/LaAlO_3$ heterostructures

As noted in the preceding section, there is an extremely high carrier density at the STO/LAO interface, and we explained this via electrons intrinsic to the polar/nonpolar interface present in the STO/LAO system. However, there are additional effects

not considered in this model. Specifically, at such high carrier densities, there is a large electric field present close to the interface in the "well" layer where the electrons reside. For $SrTiO_3$, the dielectric function is a function of field,[161] and the field in turn is a function of carrier concentration. Thus the distribution of carriers at the interface will affect the field in the $SrTiO_3$, which in turn will affect the distribution of carriers at the interface. In order to properly model this, we must self-consistently solve the Schrödinger-Poisson equation over an STO/LAO heterostructure, using first principles and experimental input. This work was performed in collaboration with Anna Sarwe, and a manuscript is in preparation.[124]

4.2.2.1 Implementation of the variable dielectric function

The Schrödinger-Poisson simulations were performed using the nextnano^3 simulation software,[10,11] although additional methods were required to implement the field-dependent dielectric function, which is not supported in nextnano^3. This allows us to solve for the electrostatic potential, charge density, and Fermi level across the STO/LAO heterostructure. The input parameters include electron effective masses (1.0 m_e, fitted to reproduce the first-principles density of states) and the bulk dielectric constant of $LaAlO_3$, 27.[162] We used a $LaAlO_3$ band gap of 4.88 eV, an $SrTiO_3$ gap of 3.27 eV, and a valence-band offset of 0.28 eV.[132] No background doping was assumed, and a two-

dimensional electron gas was inserted at the interface, with a total integrated density of 3.3x10^{-14} cm^{-2}.

In order to determine the form of the field-dependent dielectric function, we fit the results of Wolf *et al.*[161] to an equation of the following form, with a low-field cutoff at 0.2 MV/cm where ϵ_0=300:

$$\epsilon = \frac{dD}{dE} = ae^{-bE} + ce^{-fE}, \tag{4.1}$$

where ϵ is the dielectric function, D is the electric displacement, E is the applied electric field, and a, b, c and f are all fitting parameters. This allows us to determine a functional form for the field-dependent (continuous) dielectric function, and this is shown in Fig. 4.4.

However, when solving the Schrödinger-Poisson equation over a heterostructure, we cannot directly implement a continuous dielectric function. Rather, we must partition the high-field regions of the device into multiple regions with different constant dielectric functions. To determine the piecewise constant dielectric function, we integrate over E for fields greater than the cutoff field $C = 0.2$ MV/cm:

$$D = \int_0^D dD = \int_0^C \epsilon dE + \int_C^E ae^{-bE} + ce^{-fE} dE. \tag{4.2}$$

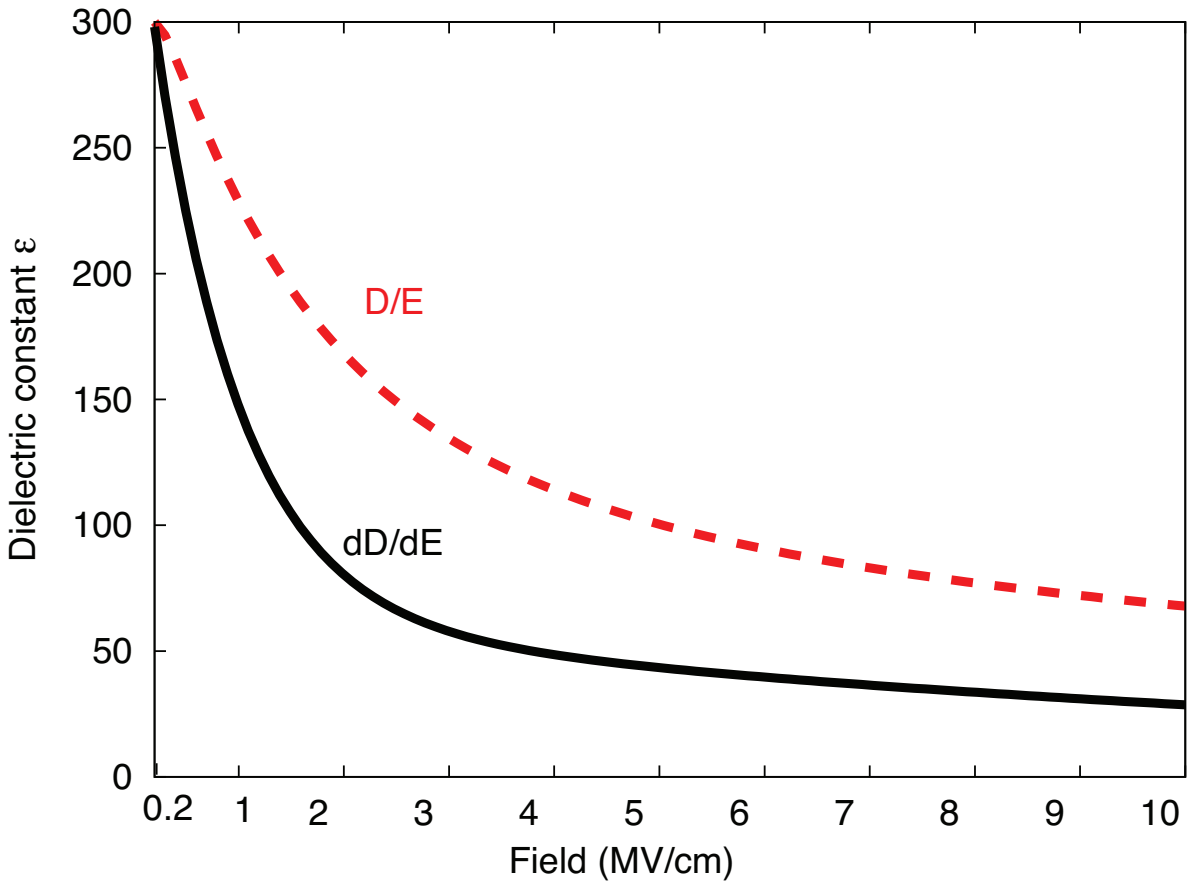

Figure 4.4: Dielectric function as a function of applied electric field. The solid black line (continuous dielectric function) is determined by fitting equation 1 to the experimental measurements,[161] with a low-field cutoff at 0.2 MV/cm. The red dashed line (piecewise constant dielectric function) is determined by integrating over E for fields greater than 0.2 MV/cm, using the fitted parameters in equation 2.

Performing this integral, and using the parameters determined from fitting to experiment allows us to determine the piecewise constant dielectric function ϵ(E), shown in Fig. 4.4:

$$\epsilon(E) = \frac{D(E)}{E}. \tag{4.3}$$

To implement this, we partition the high-field regions of the STO/LAO heterostructure into individual regions, each with their own constant dielectric function, and with sizes ranging from 0.25 nm to 1 nm, depending on the field in the particular region. The self-consistent coupled Schrödinger-Poisson simulation is run iteratively. We iterate using the following procedure: first, run the calculation with the bulk dielectric constant of 300 in every region to determine the charge density and band positions; next, determine the electric field from the output potential, and use this to define an average electric field for each region; then, use Eq. 4.3 to determine the constant dielectric function in each region; finally, restart the calculation with the new dielectric constants in each region. This process is continued until convergence is reached, and the final iteration of this process is shown in Fig. 4.5.

4.2.2.2 Impact of the variable dielectric function

We find that the electric field reaches a high of 5 MV/cm at the STO/LAO interface, which lowers the dielectric constant from 300 in the bulk to 100 under high fields. This

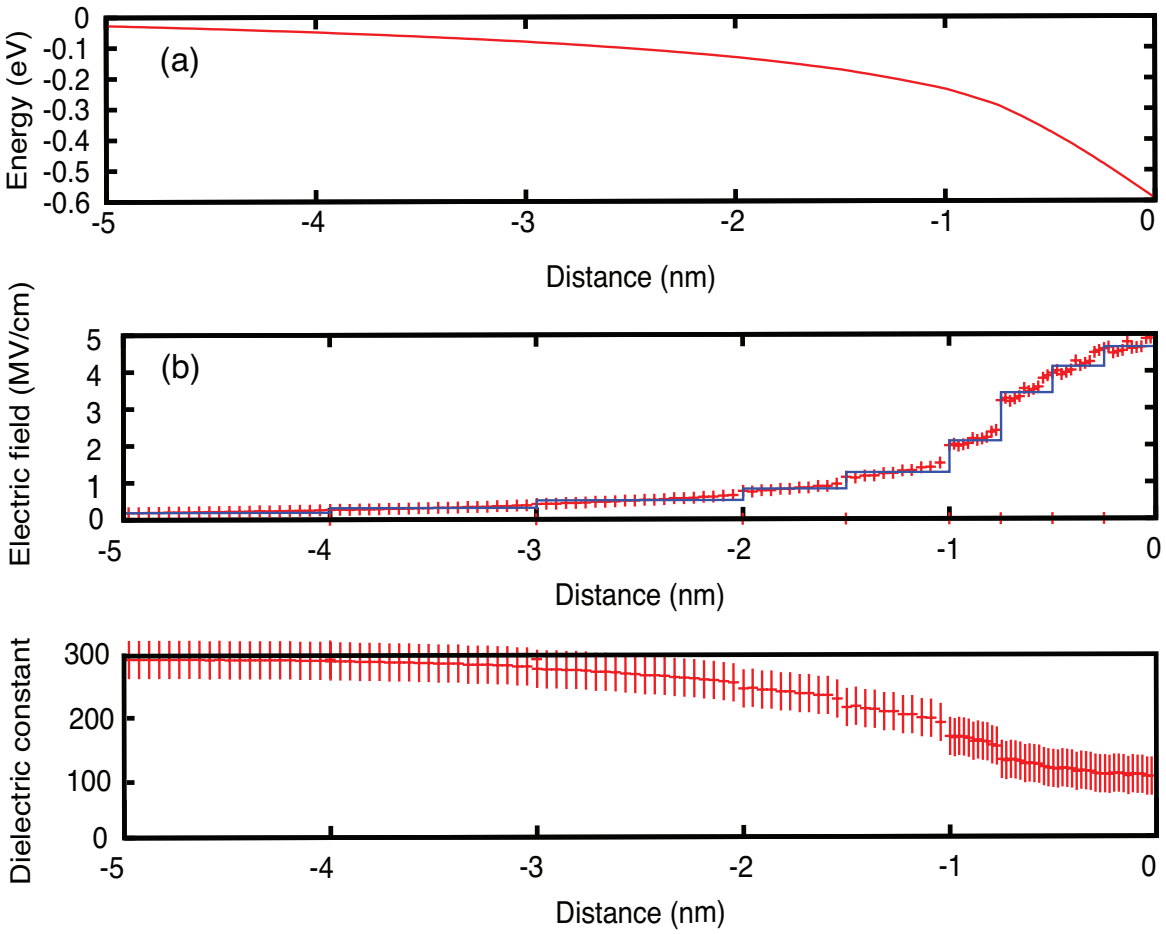

Figure 4.5: a) Fermi energy with respect to the conduction-band minimum with a variable dielectric function. b) Calculated electric field over high-field region of $SrTiO_3/LaAlO_3$ heterostructure. The stepped blue line denotes the average electric field in each cluster. c) The dielectric constant determined from the electric field.

leads to increased localization of charge carriers at the interface, with a peak in the electron distribution at just 0.5 nm away from the interface.

We can directly compare the effect on the electron distribution between the field-dependent dielectric function, and the single-valued dielectric function, shown in Fig. 4.6. In general, we expect a high density of carriers to lead to a large electric field at the interface: this in turn, will lead to a lower dielectric constant, and increased localization of carriers at the interface. This is explicitly illustrated in Fig. 4.6, where the peak of the electron distribution increases in intensity, and shifts closer to the interface: using a variable dielectric function the peak of the electron distribution is 0.5 nm from the interface, while a single-valued dielectric function results in a peak 0.8 nm from the interface.

In conclusion, we have studied the effect of a variable dielectric function on the STO/LAO interface, using fully self-consistent Schrödinger-Poisson calculations informed by first-principles and experimental input. We have found that the effect of a variable dielectric function is significant, leading to enhanced localization and increased density of charge carriers at the interface. This new understanding will help lead the way towards enhanced control over the 2DEG at the STO/LAO interface, and other complex oxide interfaces with $SrTiO_3$.

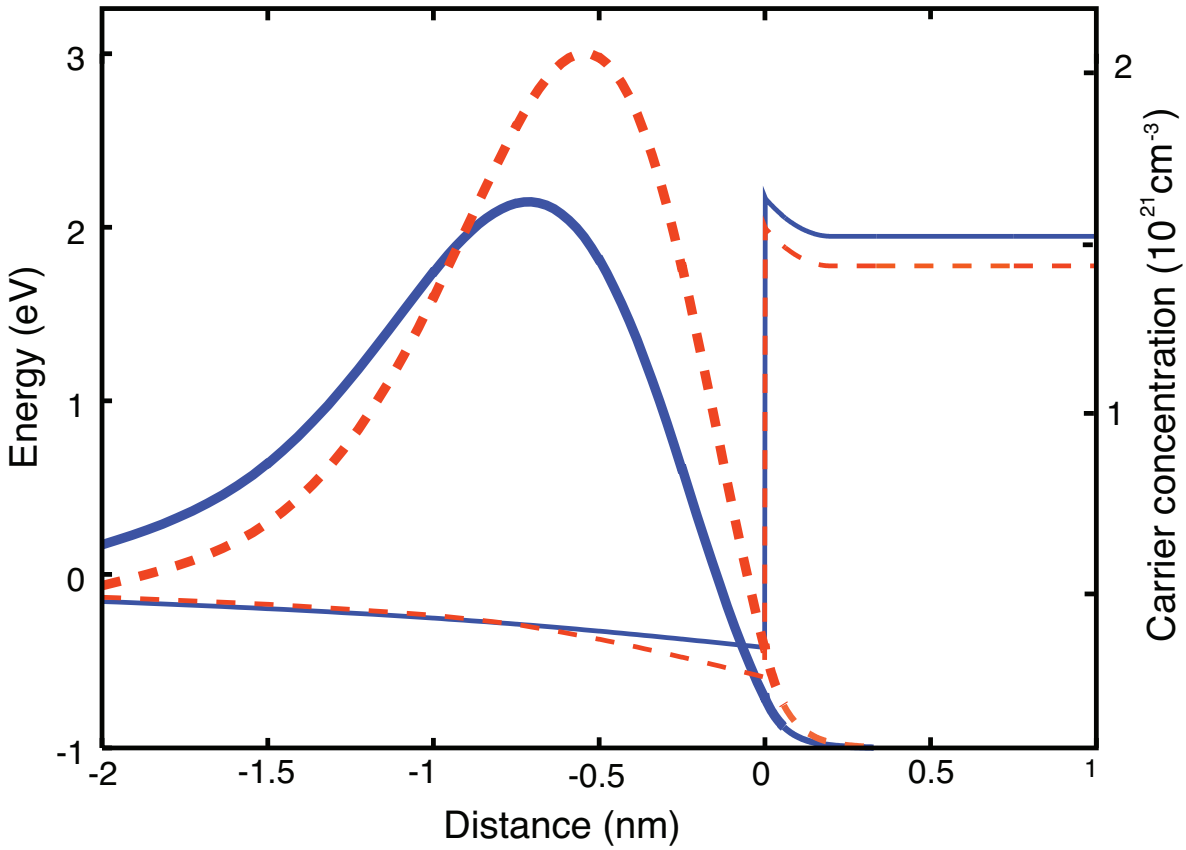

Figure 4.6: Electron distribution at the $SrTiO_3/LaAlO_3$ interface, using a variable dielectric function (red dashed line), and the bulk dielectric constant of 300 (blue solid line).

4.2.3 Alternative oxide heterostructures

In addition to the STO/LAO interface, several other oxide interface combinations have been experimentally investigated,[130,131] and many more have been predicted.[132] In particular, we investigate $BaSnO_3$ (BSO), a wide-band gap non-polar semiconductor that has the cubic perovskite crystal structure. The band structure of BSO features a highly dispersive conduction band, derived from Sn $5s$ orbitals, separated by a band gap of approximately 3 eV from a rather flat valence band, derived from O $2p$ orbitals. The exact value of the band gap, whether direct or indicted, is still a subject of debate, with values varying within 0.5 eV. First-principles calculations, performed by my collaborator Burak Himmetoglu,[125] have shown that BSO has much higher room-temperature mobility than STO, and this has been explained as a combined result of BSO's small effective mass and its small electron-phonon scattering rates. In collaboration with Karthik Krishnaswamy, and using first-principles results from Lars Bjaalie and Anderson Janotti,[132] we study the properties of a 2DEG at the STO/BSO interface using Schrödinger-Poisson simulations.

4.2.3.1 First thoughts on $BaSnO_3$

While the low effective mass of BSO reduces the DOS of conduction electrons and is therefore unfavorable for a high density 2DEG, our simulations show that one still obtains reasonably large densities at the STO/BSO interface. Given the experimental

developments leading to layer by layer growth of perovskite materials, our results indicate that BSO is a plausible candidate for the formation of a high mobility 2DEG at an interface. In addition, based on the insights we provide, we may be able to identify other materials with similar properties to BSO, that could achieve even higher 2DEG densities and lead to the development of new devices.

The small effective mass in BSO will lead to large mobility values, leading to enhanced conductivity at the interface. However, the low density of states will limit the electron density in this material, as the Fermi level (E_{F}) rises rapidly with the number of carriers in the conduction band and thereby leads to electron spillover into the STO. Fig. 4.7 shows the electron density as a function of chemical potential for BSO and STO, calculated by my collaborator Daniel Steiauf using a numerical solution of the charge-neutrality conditions. Compared to STO, which has a larger effective mass (1.9 m_e), and hence a much higher density of states, the electron density of BSO saturates much more rapidly with respect to chemical potential. For example, at an electron density of about 10^{20} cm^{-3}, E_{F} in BSO is 0.6 eV above the conduction-band minimum (CBM), whereas in STO it is only 0.1 eV above the CBM. This rapid increase in $\mathrm{E_F}$ limits the 2DEG in BSO to a maximum electron concentration, proportional to its conduction-band offset with the adjacent material .

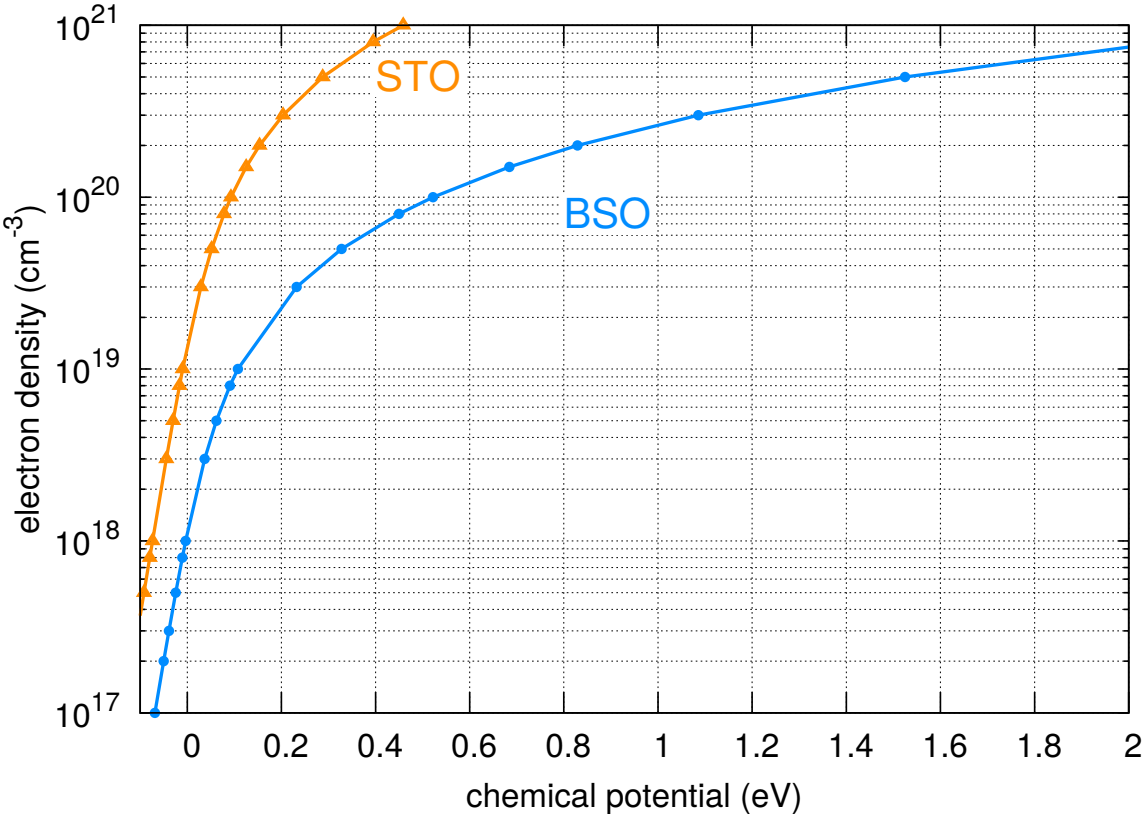

Figure 4.7: Electron density (cm^{-3}) as a function of chemical potential (eV) in BSO (solid circles) and STO (solid triangles). Calculations performed by Daniel Steiauf.

4.2.3.2 Schrödinger-Poisson simulations of $BaSnO_3$ devices

In order to more quantitatively study this system, we explicitly simulated an interface between BSO and STO, using fully self-consistent Schrödinger-Poisson simulations. The input parameters such as band gap, dielectric constant, effective mass and band offsets were extracted from first-principles calculations performed by Burak Himmetoglu and Lars Bjaalie.[125,132] A background n-type doping density of 10^{17} cm^{-3} was used for both

BSO and STO. We used modulation doping at 2 nm away from the interface at the STO side to introduce carriers near the interface. Fig. 4.8 plots the electron density (in cm^{-2}) on the BSO side of the interface as a function of the dopant density introduced through modulation doping. A complete transfer of electrons from ionized donors to the BSO region results in an "ideal" 2DEG density, as indicated by the dashed (red) line in the plot.

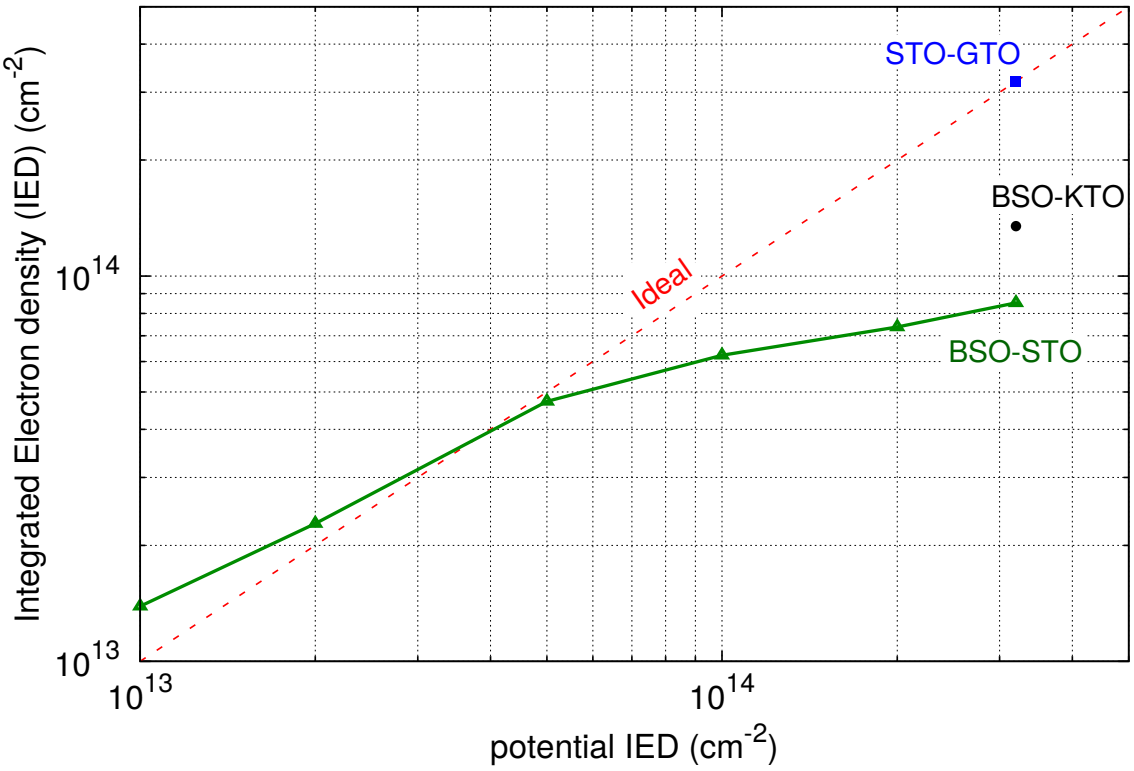

Figure 4.8: Integrated electron density (IED) confined in $BaSnO_3$, as a function of dopant density (green line). The dotted red line indicates the electron density if all dopants are converted into electrons on the $BaSnO_3$ side of the interface. Data points for $SrTiO_3/GdTiO_3$ and $BaSnO_3/KTaO_3$ are shown for comparison.

Up to a dopant density of about 5×10^{13} cm^{-2}, the 2DEG density in BSO follows the ideal behavior: i.e., all dopants directly contribute to the 2DEG density in the BSO. At dopant densities greater than 5×10^{13} cm^{-2}, only a fraction of the electrons from the dopants end up forming the 2DEG, while the rest of the electrons stay in the STO region. This is due to the Fermi level in BSO near the interface rising above the CBM of STO, leading to electron spillover. Fig. 4.9 shows the electron distribution for the BSO/STO interface with dopant density of 10^{14} cm^{-2}.

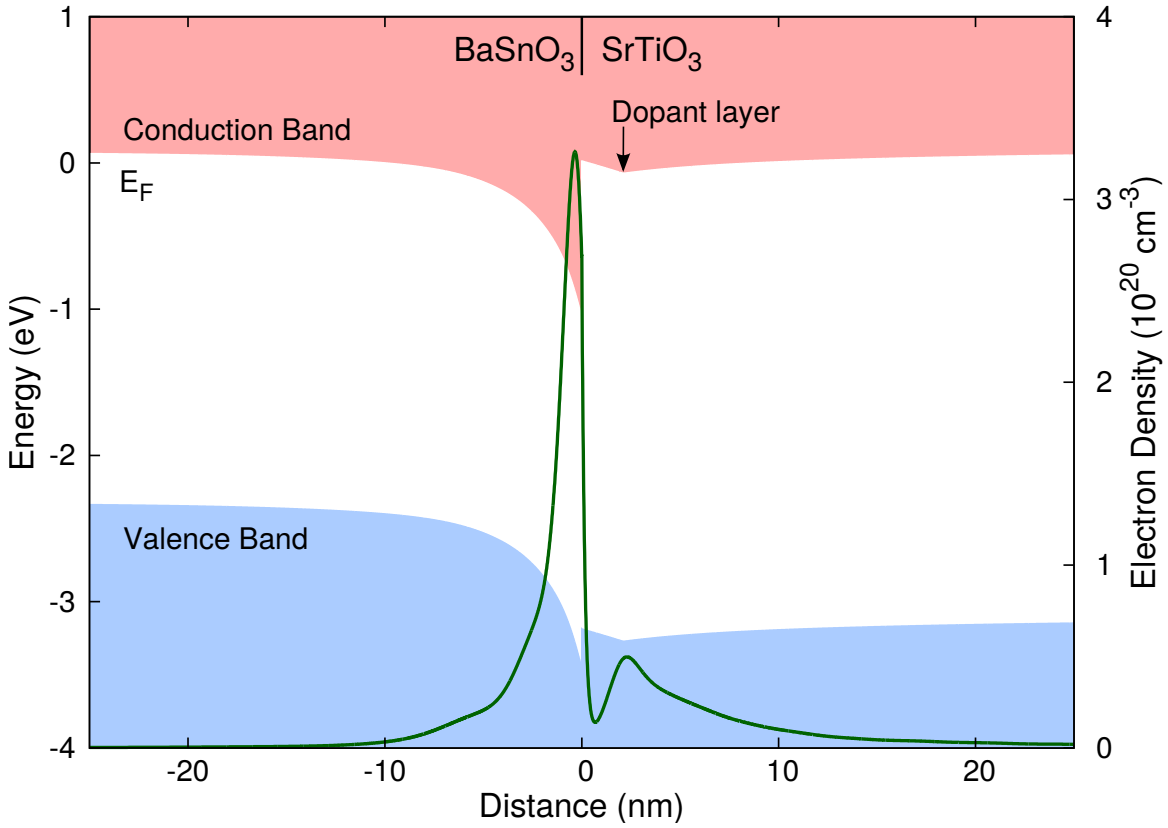

Figure 4.9: Simulated band diagram and electron density profile (solid dark green curve) of $BaSnO_3/SrTiO_3$ interface with 10^{14} cm^{-2} doping density.

In collaboration with my coauthors, I propose that a similar interface with a *polar* material may also lead to an interface with a high-density 2DEG. In particular, we identify $KTaO_3$(KTO) as a material with promising properties. Such an interface could lead to extremely high 2DEGs, on the order of 10^{14} cm^{-2}; however, the low density of states in BSO will limit the 2DEG density, as demonstrated above. To quantify this, we simulated the BSO/KTO interface with a polar discontinuity at the interface, analogous to the STO/LAO polar/non-polar interface, and with a potential 2DEG density of 3.3×10^{14} cm^{-2}. The 2DEG density confined in the BSO region was found to be about 1.5×10^{14} cm^{-2} (Fig. 4.8), which is lower than the density of 3.3×10^{14} cm^{-2} observed at the STO/GTO interface.[158]

4.3 Nitride-based devices

As discussed in Chapter 2, nitride-based devices have a range of current and potential applications, from power electronics to optoelectronic devices. In this section, we will show how macroscopic Schrödinger-Poisson simulations can be used in combination with first-principles results to model the effects of surface donor states and microscopic hole traps on the operation of III-nitride based high-electron-mobility transistors (HEMTs).

4.3.1 The impact of surface donor states on AlGaN/GaN heterojunctions

AlGaN/GaN HEMTs have recently become a subject of intense research, because of their advantages of wide band gap and high breakdown field that favor their application in high power, high frequency, and high voltage devices.[19,163] In contrast to conventional HEMTs, no intentional doping is required to generate the sheet of highly mobile electrons at the AlGaN/GaN interface.[164,165] However, a number of problems still hamper further development and integration with other devices, most notably, a lack of understanding of the origin (and ensuing lack of control) of the two-dimensional electron gas (2DEG). Among several mechanisms, the most widely accepted one is that 2DEG originates from donor states on the AlGaN surface.[166–170]

The relation between surface states and the 2DEG has been established in several experiments, including studies of the dependence of the 2DEG density on the thickness of the AlGaN barrier. When the AlGaN layer is very thin, the surface donor levels lie below the conduction-band minimum (CBM) of GaN, and no electrons are supplied to the potential well at the AlGaN/GaN interface. With increasing AlGaN barrier thickness, the polarization field leads to an increase in the surface potential until a critical point is reached where the occupied surface donor levels align with the CBM of GaN [see Fig. 4.10(a)]. Electrons can then transfer from the surface states to the

AlGaN/GaN interface, leading to 2DEG formation. If the density of occupied surface states is high, the transfer of electrons will have only a small effect on the Fermi level at the surface; as a consequence, the surface barrier height (the energy difference between the Fermi level and the CBM) of AlGaN) is independent to the AlGaN barrier thickness [Fig. 4.10(a)]. On the other hand, if the surface-state density is low and distributed over a range of energies, the transfer of electrons will lead to the depletion of the higher donor states at the surface, and a downward shift of the Fermi level will be necessary to transfer more electrons to the 2DEG. As a result, the surface barrier height will increase with the AlGaN barrier thickness [Fig. 4.10 (b)].

Experimental observations corresponding to both cases have been reported.[166,167,169–171] For example, Ibbetson *et al.*[167] found a fixed surface barrier of 1.65 eV below the CBM for $Al_{0.34}Ga_{0.66}N$. On the other hand, a change of the surface barrier height from 1.0 eV to 1.8 eV was found by scanning Kelvin probe microscopy[172] for $Al_{0.35}Ga_{0.65}N$ barriers with thickness varying from 5 nm to 44 nm. Recent experimental work by Chowdhury *et al.*[171] and by Higashiwaki *et al.*[173] measured the 2DEG density of AlGaN/GaN heterostructures at different thicknesses. These results showed large increase of surface barrier height from 1.0 to 2.5 eV when the AlGaN thickness changes from 5 to 25 nm, which would indicate that the donor states at the AlGaN surface are low in density and distributed over a large energy range.

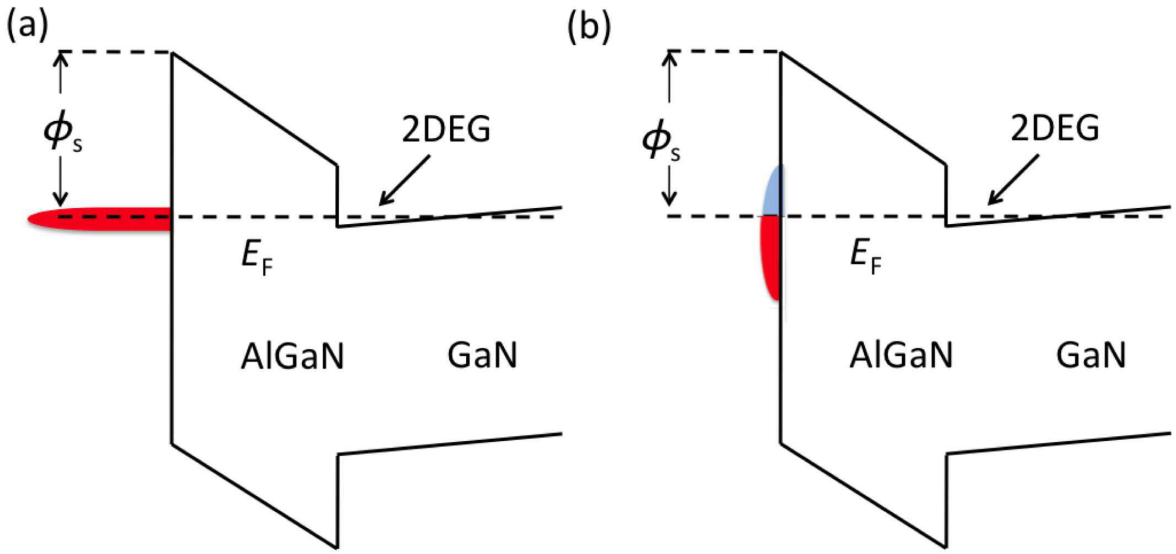

Figure 4.10: Schematic band diagrams illustrating the formation of a 2DEG with electrons supplied by surface donor states. Two cases with different distributions of the surface donor states in the energy gap are shown: (a) single surface donor level with high density; (b) distributed surface donor states with low density. The slope of the bands in the AlGaN layer is determined by the polarization discontinuity between GaN and AlGaN.

In order to study how the HEMT performance and 2DEG properties are determined by device features such as the AlGaN composition and thickness, macroscopic device simulations based on coupled Schrödinger-Poisson equations can be performed for AlGaN/GaN heterojuctions.[174–176] The surface barrier height can also be obtained by fitting simulations to experimental measurements of the 2DEG densities. However, the previous work did not explicitly include surface states in the simulations. To the extent that surface states were considered, they were represented by a single state with infinite density, which is equivalent to assuming a fixed Fermi level at the surface. Such approaches therefore cannot address the dependence of the surface barrier height on AlGaN thickness. The previous work also did not examine how the surface features, such as changes in the density of states (DOS), alter the properties of the 2DEG. Furthermore, although a number of experiments observed fixed surface barrier heights, it is unclear whether it is caused by the presence of a single (sharply peaked) surface donor state or by distributed surface donor states with high density.

In the present approach, we implement distributed surface donor states with finite density in Schrödinger-Poisson simulations and apply the simulations to an analysis of AlGaN/GaN heterojunctions. We are able to fit directly to the experimental results for 2DEG density as a function of AlGaN thickness and to extract properties of the surface such as the donor state density and the energy of the highest occupied donor state. The relation of these surface properties to the structural features of the surface is discussed.

We will also address the question of how high a donor density is needed to pin the Fermi level at the surface.

4.3.1.1 Implementation of distributed surface states

The GaN based HEMTs generally consist of an undoped GaN film with a thickness of several hundred nm up to several μm and an AlGaN layer grown pseudomorphically on GaN with a thickness ranging from a few nm to a few tens of nm. In this work, we concentrate on studying heterojunctions grown along the (0001) orientation (Ga polar HEMTs). The simulation is performed by self-consistently solving the coupled one-dimensional Schrödinger-Poisson equations along the growth direction. Within effective mass theory, the Schrödinger equation that governs the electronic states in the AlGaN and GaN regions takes the form:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dz^2}\psi(z) + (eV(z) \quad + \quad \Delta E)\psi(z) = E\psi(z), \tag{4.4}$$

where $m, \psi(z), V(z)$, and $\Delta E(z)$ are the effective mass, electron wavefunction, electrostatic potential, and the potential offset due to the conduction-band discontinuity between AlGaN and GaN, respectively. The electrostatic potential can be obtained by solving Poisson's equation:

$$\begin{aligned} \frac{d}{dz}(-\epsilon\frac{d}{dz}V(z) + P(z)) \quad &= \\ e(p(z) - n(z) \quad &+ \quad N_d^+ - N_a^-) + N_{\text{surf}}^+\delta(0), \end{aligned} \tag{4.5}$$

where ϵ is the dielectric constant, $P(z)$ is the total polarization,[177] $n(p)$ is the electron(hole) charge concentration, $N_d^+(N_a^-)$ is the ionized donor (acceptor) density and $N_{\rm surf}^+$ is the ionized donor charge at the surface. If GaN is strain free, the piezoelectric field is 0 in this region. The total polarization is then given by:

$$P(z) = P_{\rm SP}({\rm AlGaN}) + P_{\rm PE}({\rm AlGaN}) - P_{\rm SP}({\rm GaN}). \tag{4.6}$$

Instead of assuming a single surface donor level, we assume the surface donor states are continuously distributed over a fairly wide energy range with finite density and can be described by their density of states (DOS), $\rho(E)$ (see Fig. 4.11). The electron density transfered from the surface can be determined by integrating between the actual surface barrier height Φ_s and the critical surface barrier height Φ_c, which corresponds to the highest occupied surface donor state:

$$n_{\rm surf} = \int_{\Phi_c}^{\Phi_s} \rho(E)\, dE. \tag{4.7}$$

The precise features of the surface state distribution are not known. We therefore approximated it by a "rectangular" DOS, i.e., a DOS which is constant (equal to n_0) within a given energy range. We have found that the qualitative conclusions of our study are not sensitive to the details of the shape of this distribution. The corresponding density of electrons that are transferred out of the surface states is then:

$$n_{\rm surf} = n_0(\Phi_s - \Phi_c). \tag{4.8}$$

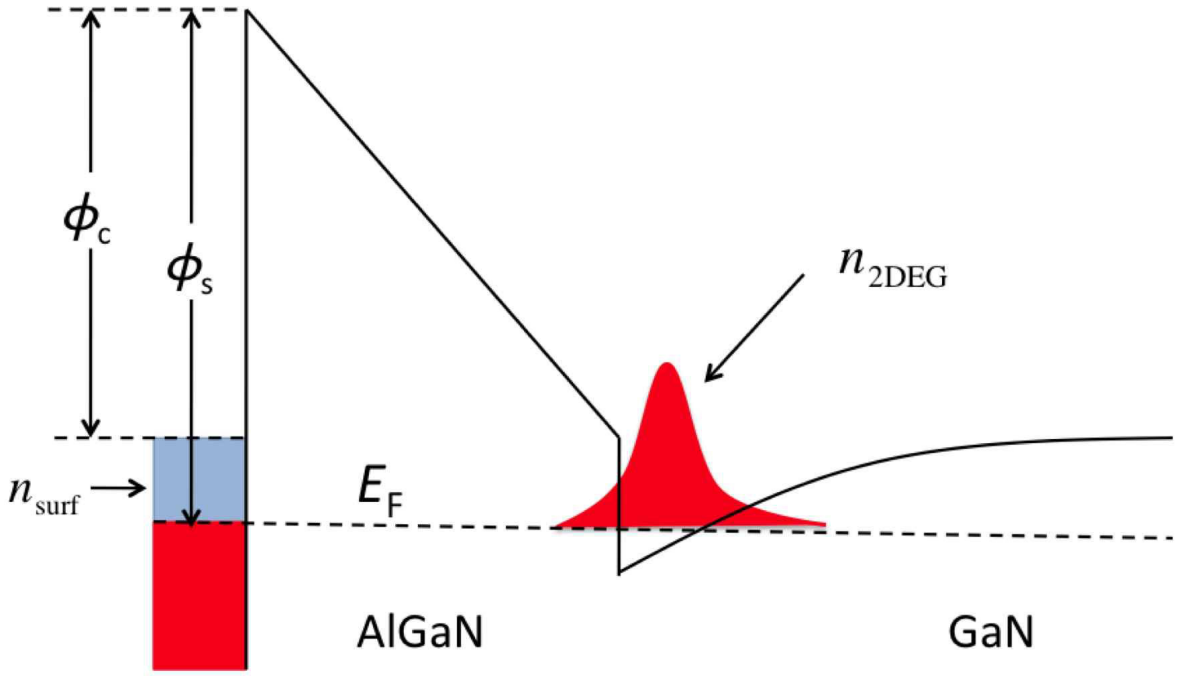

Figure 4.11: Schematic view of charge balance between 2DEG and surface charges.

On the other hand, the 2DEG charge density can be obtained by integrating the electron density around the interface. In actual devices the electron density in the AlGaN and GaN bulk region is very low, and the 2DEG charge density can be well approximated by integrating the electron density over the entire simulation region, as shown in the following equation:

$$n_{\mathrm{2DEG}} = \int_0^L n(z)dz, \tag{4.9}$$

in which 0 represents the AlGaN surface and L is the total length of the AlGaN and GaN regions.

The samples used by my collaborators in Ref. 126 are not intentionally doped. The unintentional doping level is quite low, on the order of 10^{16} cm^{-3}. If the depletion length is 100 nm, the 2DEG density contributed by ionized donors is only 10^{11} cm^{-2}, which is insignificant compared to the contribution from surface donor states. Therefore we assume in the simulations that *all* the electrons in the 2DEG are contributed by surface states, i.e., $n_{\mathrm{2DEG}} = n_{\mathrm{surf}}$, and devise an iterative approach for determining the electron density due to the surface donor states. In the first simulation step, an initial value is given to Φ_s. The 2DEG density, $n_{\mathrm{2DEG}} = \int n(z)dz$, and n_{surf} are calculated using the current Φ_s. The two densities are then compared. If n_{2DEG} is larger than n_{surf}, the surface barrier height will be increased; otherwise, it will be decreased. This is repeated until n_{2DEG} equals n_{surf}. The resulting n_{2DEG} and Φ_s are the self-consistent 2DEG charge

density and surface barrier height. This simulation procedure has been implemented in the semiconductor device simulator nextnano^3.[10,11]

The simulations are performed for the (0001) orientation, assuming pseudomorphic growth on a GaN substrate, i.e., there is no strain on the GaN region, whereas the strain in the AlGaN region is calculated from the difference in lattice parameters between AlGaN and GaN. A background doping level of 1.6×10^{16} cm^{-3} is uniformly distributed throughout both the AlGaN and GaN regions. This low doping density represents the unintentional doping in the layers. The thickness of GaN is 300 nm, whereas the thickness of AlGaN various from 5 nm to 25 nm. Table 4.1 lists the values of all relevant parameters used in the simulations, including lattice constants, piezoelectric constants, spontaneous polarization as well as the (relative) band positions for AlN and GaN. Properties of AlGaN alloys are obtained by linear interpolation, taking bowing in the AlGaN band gap into account.

4.3.1.2 Fixed versus variable surface barrier heights

Before attempting to explicitly simulate the experimental results, we first examine how the density of surface donor states affects the 2DEG density and its dependence on AlGaN thickness. We first perform simulations using single surface donor levels with energies of 1 eV, 1.65 eV, 2 eV, 3 eV and 4 eV below the AlGaN CBM, for an Al composition of 34%. As shown in Fig. 4.12 (dotted lines), the 2DEG electron density increases

Table 4.1: Parameters used in our simulations, as taken from Refs. 178,179. The VBM energy of AlN is set to 0 and the energies of the AlN CBM as well as the GaN VBM and CBM are given by the AlN and GaN band gaps and their band offsets. A valence-band offset of 0.8 eV is used.

		x	y	z
AlN				
Lattice constant (Å)	-	3.112	3.112	4.982
Piezoelectric constants (Cm^{-1})	-	1.800	-0.640	-0.480
Spontaneous polarization (Cm^{-1})	-	0.000	0.000	-0.100
Elastic constant C_{33} (GPa)	373			
C_{31}(GPa)	108			
Dielectric constant	9.6			
CBM position (eV)	6.238			
VBM position (eV)	0			
GaN				
Lattice constants (Å)	-	3.189	3.189	5.185
Piezoelectric constants (Cm^{-1})	-	0.860	-0.440	-0.300
Spontaneous polarization (Cm^{-1})	-	0.000	0.000	-0.032
Elastic constant C_{33} (GPa)	398			
C_{31}(GPa)	106			
Dielectric constant	9.8			
CBM position (eV)	4.315			
VBM position(eV)	0.8			

rapidly once the AlGaN thickness exceeds a critical value that is determined by the energy of the donor level. At large thickness values, the 2DEG will completely screen the polarization field. Therefore n_{2DEG} approaches the polarization charge asymptomatically at large AlGaN thicknesses, irrespective of the position of the surface donor level. However, the AlGaN thickness in HEMT devices is generally limited to a few tens of nanometers. In this thickness range, the 2DEG densities are considerably different for different surface barrier heights. For example, at an AlGaN thickness of 10 nm, the n_{2DEG} values are 1.32×10^{13} cm^{-2} and 0.63×10^{13} cm^{-2}, for surface donor levels at 1.65 eV and 3.0 eV, respectively.

Next, simulations for distributed surface donor state densities are performed assuming a critical surface barrier height of 1.65 eV and using different n_0 values, including 0.5×10^{13} cm^{-2}eV^{-1}, 1.0×10^{13} cm^{-2}eV^{-1} and 1.0×10^{14} cm^{-2}eV^{-1}. In this case, our simulations demonstrate that the surface barrier height will increase with increasing AlGaN thickness. For instance, as shown in Fig. 4.12, for $n_0 = 1.0 \times 10^{13}$ cm^{-2}eV^{-1} the 2DEG density is significantly lower than the one obtained from the simulation with a surface donor level fixed at 1.65 eV. We note that at an AlGaN thickness of 30 nm, n_{2DEG} is very close to the electron density obtained assuming a fixed surface donor level at 3.0 eV, indicating that the surface barrier height has increased from 1.65 eV to about 3.0 eV. The simulations in Fig. 4.12 show that the changes of Φ_s will be larger if n_0 is smaller. For example, Φ_s changes from 1.65 eV to about 4 eV for an n_0 value of

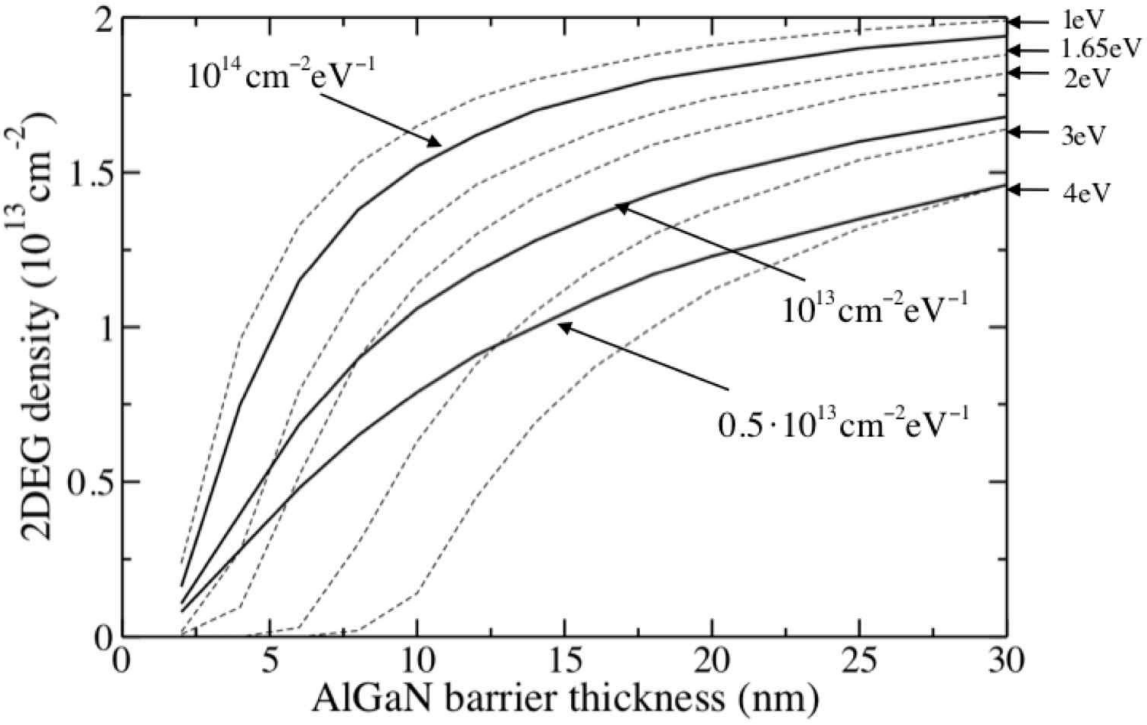

Figure 4.12: Simulated 2DEG densities as a function of AlGaN barrier thickness for $Al_{0.34}Ga_{0.66}N$ with fixed (dotted lines) and variable (solid lines) surface barrier heights. For the latter, a critical surface barrier height Φ_c of 1.65 eV was assumed.

0.5×10^{13} cm^{-2}eV^{-1}. On the other hand, for large values of n_0 the changes in surface barrier height become small. As shown in Fig. 4.12, for n_0=n_0 = 10^{14} cm^{-2}eV^{-1}, the simulated n_{2DEG} values are quite close to those obtained from a simulation for fixed surface barrier height. This result indicates that although some experiments (e.g., Ibbetson *et al.*[167]) reported a fixed surface barrier height, their results do not necessarily correspond to a single surface donor level. In theory, the density of a single surface donor level on the AlGaN (0001) surface could be as high as 1×10^{15} cm^{-2}.[180,181] As we just saw, the surface barrier height will appear to be fixed when the surface state density is on the order of 1×10^{14} cm^{-2}eV^{-1} or higher.

4.3.1.3 Comparing with experiment: the reality check

In experimental work by Chowdhury *et al.*[171] and by Higashiwaki *et al.*,[173] the 2DEG densities as a function of AlGaN thickness were obtained by Hall measurements for three Al compositions, 19%, 24% and 29%. Using the simulation method and the HEMT model described earlier, we now attempt to reproduce these results. The simulation results are presented and compared with experiments in Fig. 4.13 (a)-(c). For each composition, the density n_0 and the critical surface barrier height Φ_c were optimized by fitting to the experimental results. The corresponding values are presented in Table 4.2, along with the values reported in a previous pubblication which were obtained by fitting each experimental point by a simulation with fixed surface donor level.[173]

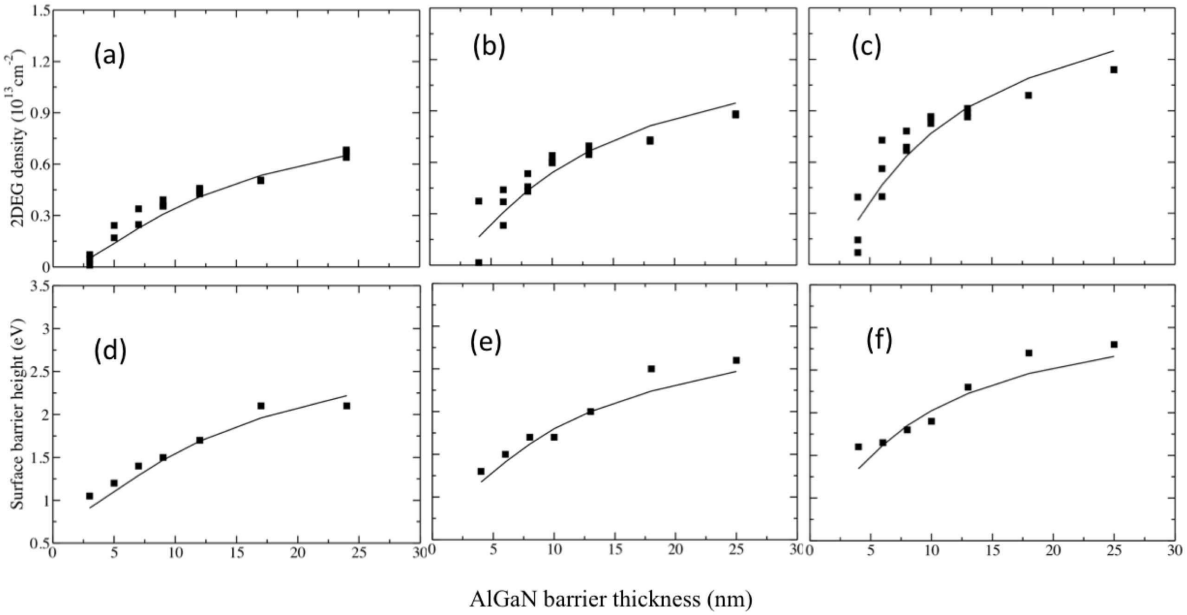

Figure 4.13: Simulated 2DEG densities [(a)-(c)] and surface barrier heights [(d)-(f)] as a function of AlGaN barrier thickness for Al compositions of 19%, 24% and 29%. The simulated results are shown in solid lines; experimental results from Ref. 126 are indicated by filled squares. Measurements were performed on multiple samples for a few different AlGaN thickness values. In the case of 2DEG densities the actual experimental results from Ref. 126 are shown, but for the surface barrier height only averaged values are displayed.

Table 4.2: Parameters describing surface donor states obtained by fitting 2DEG densities to experimental results. n_0 characterizes our assumed rectangular DOS values and Φ_c is the critical surface barrier height (corresponding to the highest occupied surface state). The change $\Delta\Phi$ of the surface barrier height from its critical value to its value at an AlGaN thickness of 25 nm are is also listed, and compared with the$\Delta\Phi_{\mathrm{Old}}$ value obtained in previous work from simulations using single surface donor levels.

Material	$n_0(\mathrm{cm}^{-2}\ \mathrm{eV}^{-1})$	Φ_c(eV)	$\Delta\Phi$ (eV)	$\Delta\Phi_{\mathrm{Old}}$[173]
$Al_{0.19}Ga_{0.81}N$	0.46×10^{13}	0.8	1.31	1.1
$Al_{0.24}Ga_{0.76}N$	0.61×10^{13}	0.9	1.30	1.32
$Al_{0.29}Ga_{0.71}N$	0.75×10^{13}	1.0	1.32	1.23

The critical surface barrier height Φ_c corresponds to the highest occupied donor levels on the surface. The fitted values are close to 1 eV and show a slight increase with increasing Al composition. All of these values are significantly lower than the value of 1.65 eV reported for a single donor level on $Al_{0.34}Ga_{0.66}N$ by Ibbetson *et al.*,[167] but in good agreement with Koley *et al.*[170] who also observed a variable surface barrier height.

Fig. 4.13 (d)-(f) show the variation of the surface barrier height with AlGaN layer thickness, compared with the values obtained in previous work by simulations using fixed surface donor levels.[171,173] The changes of the surface barrier height from an AlGaN thickness of 3 nm (19%) or 4 nm (24% and 29%) to 25 nm are also listed in Table 4.2. As shown in Fig. 4.13 and Table 4.2, the changes of the surface barrier heights are as large as 1.3 eV for the three compositions.

Overall, the simulations manage to reproduce the experimental results quite well. Still, noticeable deviations occur, especially for the samples with higher Al composition. We can speculate on several possible causes for these discrepancies, on both the simulation and the experimental side. The large strain in AlGaN due to the pseudomorphic growth on GaN may be partially relaxed for thick samples, which would lower the piezoelectric polarization and therefore the 2DEG density. This effect would be more pronounced for samples with larger Al compositions, consistent with the trends in Fig. 4.13 (a)-(c). Another issue may be that the chosen form of the surface DOS may affect the simulation results. If, instead of being constant, the density of surface donor states is higher at higher energies, the resulting n_{2DEG} would be lower for larger AlGaN thicknesses, providing a better fit to the experimental results. Finally, we note that different choices of simulation parameters such as the spontaneous polarization and the piezoelectric constants (or nonlinear variations of these quantities as a function of alloy composition) might also affect the results.

4.3.1.4 But what does it mean?

Irrespective of whether the surface barrier heights are considered variable or fixed, all experimental measurements of 2DEG densities from Ref. 126 detected the presence of surface donor states that are close in energy to the AlGaN CBM, although the origin of the surface donor levels has been ill understood.[167,171,172,182,183]

The critical surface barrier height Φ_c corresponds to the highest occupied surface donor state, and its value may be used to correlate the experimental observations with microscopic features of the surface as calculated from first principles. It is therefore worthwhile to critically examine the fitting of Φ_c to experimental results. Fig. 4.14 shows simulated 2DEG densities using Φ_c values of 1.0 eV and 1.6 eV for an Al composition of 29%. Both fits capture the general trend of the experimental results, but the one using the smaller Φ_c value fits significantly better to the points in the range from 5 nm to 15 nm, which are the more reliable ones among the experimental data. As noted above, larger deviations may occur for the 25 nm sample because of partial strain relaxation. Similarly, the best fitting values for 19% and 24% samples were found to be 0.8 and 0.9 eV. We therefore concluded that the highest surface donor state is about 1.0 eV below the AlGaN CBM, and shifts downward slightly with increasing Al concentration. Koley *et al.* who also observed a variable surface barrier height suggested a similar value (∼1.0 eV below the CBM) for the highest surface donor state.[170]

The Φ_c value in the case of a variable surface barrier height is therefore considerably smaller than the values reported in experiments consistent with fixed surface barrier height, around 1.7 eV for Al compositions around 35%.[167–169] These results suggest that differences in sample processing and consequently in surface oxidation not only lead to different *densities* of the surface donor states but also to a different *distribution* of such states in the band gap, a hypothesis supported by first-principles calculations performed

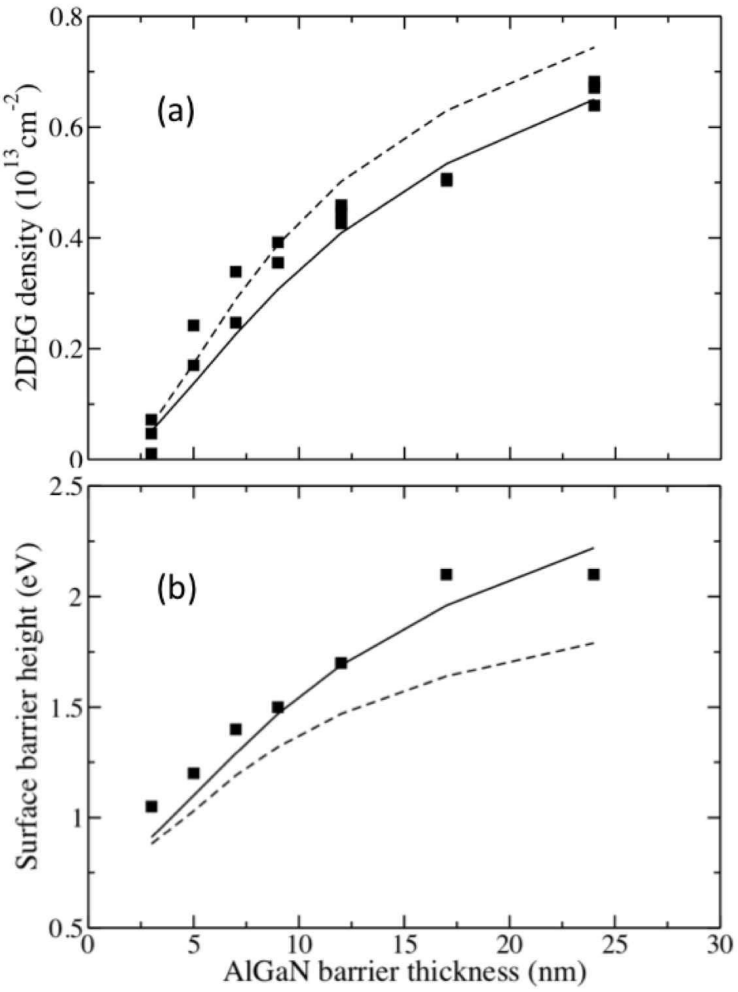

Figure 4.14: Comparison of (a) 2DEG densities and (b) surface barrier heights as a function of AlGaN barrier thickness for $Al_{0.19}Ga_{0.81}N$, simulated using two different surface DOS values; solid lines correspond to $n_0 = 0.46 \times 10^{13}$ cm^{-2}eV^{-1}, the dashed lines to $n_0 = 0.75 \times 10^{13}$ cm^{-2}eV^{-1}. The filled squares show the experimental results.

by my collaborator, Maosheng Miao,[183] which showed that the highest occupied surface donor states occur at lower energy when their density is high, explaining the experimental observations of different minimum surface barrier heights in the cases of "fixed" versus variable barrier heights.

Furthermore, we find that the fitted surface DOS values n_0 increase with increasing Al composition (Table 4.2). For an Al composition of 29%, n_0 is found to be about 60% larger than for the 19% Al composition. To examine the sensitivity of our fits to this parameter, we performed simulations of the 2DEG density and the surface barrier height for an Al composition of 19% using either $n_0 = 0.46 \times 10^{13}$ cm^{-2}eV^{-1} or $n_0 = 0.75 \times 10^{13}$ cm^{-2}eV^{-1}, the latter being the value that best fits the 2DEG density of the sample with 29% Al concentration. As shown in Fig. 4.14, the simulation using $n_0 = 0.75 \times 10^{13}$ cm^{-2}eV^{-1} provides a poor fit, particularly for the surface barrier heights. Note that a larger surface DOS value results in less change in surface barrier height for a given amount of charge transfer from the surface to the 2DEG. The increase in surface DOS for samples with higher Al composition is also consistent with the increase in critical surface barrier height (see Table 4.2), for the reasons explained above. At present, we have no direct explanation for the trend that higher Al compositions would result in a higher density of donor states. Experimental surface science studies that might show differences in surface features as a function of Al composition have not yet been performed. The higher surface donor state density may be related to a higher tendency

for forming oxide-stoichiometry-matching surface reconstructions when the Al content is higher,[126] possibly driven by the high tendency of Al to oxidize.

In collaboration with Maosheng Miao, we developed a simulation approach for AlGaN/GaN HEMTs that takes into account the contribution of surface donor states with a distributed and finite DOS in the band gap. We applied this method to simulating the dependence of 2DEG density on AlGaN barrier thickness at three different Al compositions. The fitting parameters are the magnitude of the value of the DOS (assumed to be constant) and the critical energy corresponding to the highest occupied surface state. Our results also show that the surface DOS values and critical barrier height increase with Al composition. Good agreement was obtained with experimental results obtained by my collaborators,[126] and a physical interpretation of the resulting parameters was provided.

4.3.2 Hole-trapping behavior in nitride semiconductors

Many recent nitride semiconductor devices employ heterostructures in which the Fermi level is established near the valence-band maximum (VBM) due to the influence of polarization fields. One class of devices in which this occurs are the polarization-hole-doped structures, where a polarization-induced hole gas is engineered to arise in a graded AlGaN alloy.[184] Another example is N-face HEMTs; in this case, a two-dimensional hole gas (2DHG) should be present at the interface between the AlGaN back barrier

and GaN buffer due to the polarization field.[185] In these types of heterostructures, the presence of hole traps has been found to adversely affect device performance. In the polarization-hole-doped heterostructures, such hole traps are thought to prevent the formation of field-ionized hole gases.[184] In N-face HEMTs, these hole traps are thought to cause large-signal dispersion that is observed in current-voltage curves.[185,186] In both cases, hole traps have been blamed for compensating the polarization-induced hole gases which should be present within these heterostructures.

4.3.2.1 Identification of the hole traps

Based on deep level transient spectroscopy (DLTS), it has been proposed that a deep donor defect with a transition level 60 meV from the valence-band maximum (VBM) of GaN is responsible for this hole-trapping behavior.[185,186] The hole-trapping effect can be mitigated through Si-doping of the GaN buffer,[186] which establishes the Fermi level far away from the VBM and likely prevents hole-trapping from occurring. However, the microscopic mechanism whereby such hole-trapping occurs has not yet been established.

Hybrid functional calculations were performed by my collaborator John Lyons[127] to examine possible hole-trapping mechanisms. These first-principles calculations show that hole self-trapping *does not* occur in GaN, AlN, or at the AlN/GaN interface; however, hole trapping does occur in the presence of acceptor impurities, in particular the C impurity. Using Schödinger-Poisson simulations we examine the effect of C hole traps

on N-face HEMTs, and compare this behavior with that of the deep-donor interface traps that have been proposed previously in the literature. Both the dispersed C hole trap and the interface trap are found to deplete the hole gas at the AlGaN/GaN back barrier interface, indicating that C impurities can indeed cause hole trapping in nitride heterostructures.

C is a common unintentional dopant in GaN and AlN, and is also intentionally doped into GaN to create semi-insulating material.[187] In recent work, my collaborator John Lyons has found that C_N acceptors feature (+/0) deep donor transition levels in both GaN (where C_N has a 0.35 eV ionization energy) and AlN (where C_N has a 1.07 eV ionization energy).[24] Those calculations also showed that C_N is the most stable form of the C impurity in both materials under *n*-type conditions; in addition, the C_N-H_i complex features a (+/0) transition level in GaN, 0.25 eV above the VBM, indicating that C impurities could continue to act as hole-trapping centers even if they are electrically inactive. While the hole-trapping transition levels of C are higher than the interface traps often discussed in the literature, in the following section we demonstrate that both types of traps exhibit very similar behavior in nitride heterostructures.

4.3.2.2 Implications for nitride heterostructures

We examine the effect that C hole traps could have on nitride devices by performing Schrödinger-Poisson simulations. We use the N-face HEMT device structure shown

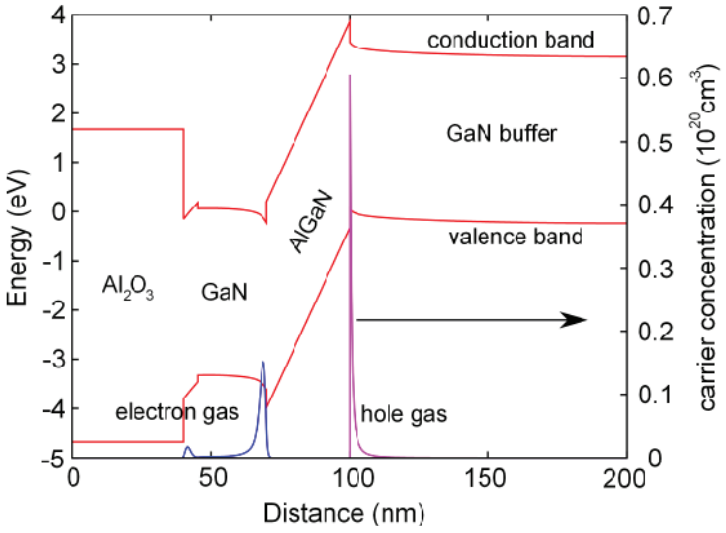

Figure 4.15: Band structure of a simulated N-face HEMT, based off of the design in Ref. 186, in which a 2DHG accumulates at the AlGaN/GaN buffer interface (located at 100 nm along the x-axis).

in Fig. 4.15 (similar to the one investigated in Ref. 186) for these simulations, which contains 40 nm of Al_2O_3 oxide, a 20 nm GaN channel, 30 nm of $Al_{0.3}Ga_{0.7}$ back barrier, and a 100 nm GaN buffer layer. Under zero bias conditions this device exhibits a two-dimensional electron gas (2DEG) at the first GaN/AlGaN interface and a two-dimensional hole gas (2DHG) at the second GaN/AlGaN interface where the valence bands of AlGaN and GaN approach the Fermi level (at zero on the left y-axis).

To investigate hole-trapping behavior in this HEMT we examine two cases. In Case 1 we place 0.5-1.5 x 10^{12} cm^{-2} deep donor traps (with a (+/0) transition level 60 meV from the VBM) at the second AlGaN/GaN interface. Both the ionization energy and concentration of this trap are similar to that proposed previously in the literature.[185,186]

In Case 2 we examine 0.5-1.5 x 10^{17} cm^{-3} C hole traps (with a 350 meV deep donor transition level) dispersed throughout the GaN buffer region.

In Figs. 4.16 and 4.17 the density of the 2DHG as a function of applied negative gate bias is shown for both trap cases. In both cases, these deep donor traps reduce the density of the 2DHG: instead of polarization charges being compensated by free holes, holes are trapped at the deep donor states, and these positively charged centers compensate the polarization charge. In fact, the C buffer traps (shown in Fig. 4.17) give rise to an even greater reduction in 2DHG than the 60 meV interface traps, indicating that these impurities could cause hole trapping, and possibly explaining the diminished output conductance in N-face nitride-based HEMTs.

We have used Schrödinger-Poisson simulations to investigate the implication of different potential microscopic mechanisms for hole trapping behavior in nitride heterostructures. Hybrid-functional calculations[127] performed by my collaborator, John Lyons, indicate that holes can trap at acceptor impurity sites, and C acceptors are therefore identified as a plausible candidate for the hole-trapping behavior. Using Schrödinger-Poisson simulations we have found that bulk C impurities lead to similar effects in transistors as we would expect from the interface hole traps that have often been proposed in the literature. Thus, we propose that unintentionally incorporated C, dispersed throughout the GaN buffer region, can be responsible for hole-trapping behavior in N-face HEMTs.

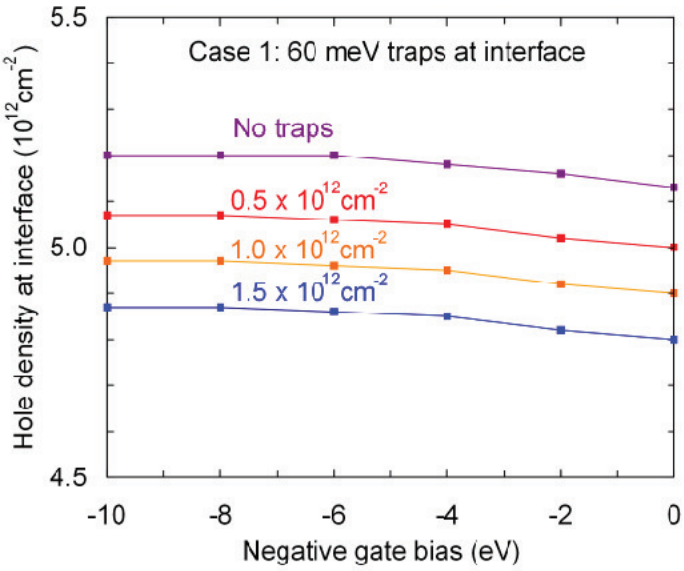

Figure 4.16: 2DHG density as a function of negative bias for the case of the 60 meV interface hole trap.

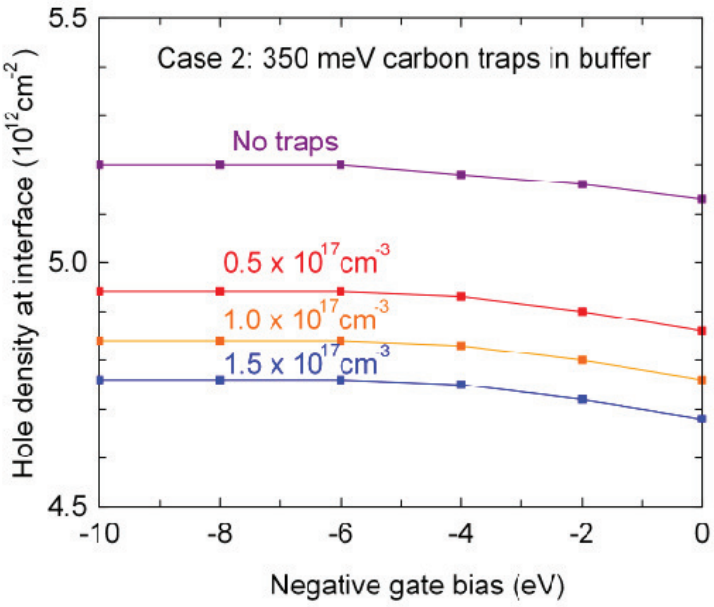

Figure 4.17: 2DHG density as a function of negative bias for the case of 350 meV C traps in the GaN buffer.

4.4 Conclusions

In this chapter, we have illustrated the power of macroscopic simulations performed in concert with first-principles input to explain and predict device performance. Specifically, we have presented calculations for oxide and nitride heterostructures. We have explained the origin of the two-dimensional electron gas (2DEG) as arising from the polar discontinuity at the STO/LAO interface, and have emphasized the need to prevent electrons draining away from the interface by preventing an electrostatic potential buildup in the LAO layer. In addition, we have studied the effect of a field-dependent dielectric function at the STO/LAO interface, and shown that it leads to an increased localization of the 2DEG at the interface. We have also studied the relationship between surface donor states and the two-dimensional electron gas in AlGaN/GaN HEMTs, and shown that the Schottky barrier in such devices is variable and highly surface state dependent. Finally, we have studied the effect of deep carbon acceptors in N-face AlGaN/GaN HEMTs, and proposed that such unintentionally incorporated C can explain the hole-trapping behavior in N-face HEMTs.

Chapter 5

Summary and future directions

> Look closely at the present you are constructing. It should look like the future you are dreaming.
>
> Alice Walker

Over the last 50 years, semiconductor devices have shrunk to the level where atomic-scale effects play a major role in determining the operation of the device. Continued improvement in device performance will no longer be possible without an in-depth knowledge of atomic physics and its implications for device performance.

In order to illustrate the power of an atomistic approach to understanding devices, we have presented results and conclusions from three interlinked projects: n-type doping of

III-nitride semiconductors, defects for quantum computing, and macroscopic simulations of devices.

First, we studied n-type doping of III-nitride semiconductors and their alloys, and analyzed the barriers to effective n-type doping of III-nitrides and their alloys. We have studied the formation of DX centers, specifically O, Si and Ge, and we predicted alloy composition onsets for various III-nitride alloys. In addition, we performed a comprehensive study of alternative dopants, and proposed sulfur or selenium as potential alternative dopants to improve the n-type conductivity in AlN and wide-band-gap nitride alloys.

Next, we discussed atomic-scale defects as a source of decoherence in superconducting qubits; we investigated the effect of two-level state defects (TLS) systems in alumina as a source of decoherence in superconducting qubits based on Josephson junctions, and showed that hydrogen incorporating in a double well configuration was responsible for this TLS decoherence. We also discussed ways by which new defects as qubits can be identified in new materials, and their spin states, formation energies, and emission spectra characterized, specifically showing calculations on various potential qubits in SiC.

Finally, we provided examples of recent calculations we have performed for devices using macrosopic device simulations, largely in conjunction with first-principles calculations. We discussed the power of using a multi-scale approach to accurately model

oxide and nitride-based heterostructures in a range of different contexts, from control and comprehension of the two-dimensional electron gas at the STO/LAO interface, to a better understanding of the effects of surface donor states and hole trapping mechanisms on AlGaN/GaN HEMTs.

Overall, we have shown that an atomistic approach to understanding device problems, combined with macroscopic device simulations and informed by experimental results, has become an increasingly necessary and effective tool for understanding device performance. As devices continue to shrink in size, the benefits of performing atomic-scale first-principles studies will increase, while the computational cost of the calculations decreases.

The original atomists were philosophers, and their abstract concept of atoms was as far removed from everyday reality as was their abstract concept of a soul. Today, the abstract has become concrete, and atoms have invaded our everyday technology. The least we can do is try to understand them.

Bibliography

[1] P. Hohenberg and W. Kohn.. *Phys. Rev.* **136** B864 (1964).

[2] W. Kohn and L. J. Sham. *Phys. Rev.* **140** A1133 (1965).

[3] J. Heyd., G. E. Scuseria., and M. Ernzerhof. *J. Chem. Phys.* **21** 219906 (2006).

[4] C. E. Dreyer. *Thesis* (2014).

[5] J. L. Lyons, A. Janotti, and C. G. Van de Walle. *Appl. Phys. Lett.* **97** 152108 (2010).

[6] P. G. Moses, M. Miao, Q. Yan, and C. G. Van de Walle. *J. Chem. Phys.*, **134** 084703 (2011).

[7] C. Freysoldt, B. Grabowski, T. Hickel, G. Kresse, A. Janotti, J. Neugebauer, and C. G. Van de Walle. *Rev. Mod. Phys.* **86** 253 (2014).

[8] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle. *Phys. Rev. Lett.* **102** 016402 (2010).

[9] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle. *Phys. Status Solidi (b)* **248** 1067 (2011).

[10] S. Birner *et al*. *Acta Phys. Pol. A* **110** 111 (2006).

[11] nextnano.de/.

[12] M. A. Nielsen and I. L. Chung. *Quantum Computation and Quantum Information*. Cambridge University Press (2000).

[13] G. Kresse and J. Furthmüller. *Phys. Rev. B* **54** 11169 (1996).

[14] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.* **77** 3865 (1996).

[15] P. E. Blöchl. *Phys. Rev. B* **50** 17953 (1994).

[16] A. Baldereschi. *Phys. Rev. B*, **7** 5212 (1973).

[17] D. J. Chadi and M. Cohen. *Phys. Rev. B*, **12** 5747 (1973).

[18] H. J. Monkhorst and J. D. Pack. *Phys. Rev. B*, **12** 5188 (1976).

[19] U. K. Mishra, P. Parikh, and Y. F. Wu. *Proc. IEEE* **90** 1022.

[20] S. Nikishin, B. Borisov, M. Pandikunta, R. Dahal J. Y. Lin, H. X. Jiang, H. Harris, and M. Holtz. *Appl. Phys. Lett.* **83** 5163 (2009).

[21] M. Bickermann, B. M. Epelbaum, O. Filip, P. Heimann, M. Feneberg, S. Nagata, and A. Winnaker. *Phys. Status Solidi. (c)* **7** 1743 (2010).

[22] N. A. Fichtenbaum, T. E. Mates, S. Keller, S. P. DenBaars, and U. K. Mishra. *J. Cryst. Growth* **310** 1124 (2008).

[23] M. Sumiya, K. Yoshimura, K. Ohtsuka, and S. Fuke. *Appl. Phys. Lett.* **76** 2098 (2000).

[24] J. L. Lyons, A. Janotti, and C. G. Van de Walle. *Phys. Rev. B* **89** 035204 (2014).

[25] C. G. Van de Walle. *Phys. Rev. B* **57** R2033 (1998).

[26] M. D. McCluskey, N. M. Johnson, C. G. Van de Walle, D. P. Bour, M. Kneissl, and W. Walukiewicz. *Phys. Rev. Lett.* **80** 4008 (1998).

[27] L. Gordon, J. L. Lyons, A. Janotti, and C. G. Van de Walle. *Phys. Rev. B* **89** 085204 (2014).

[28] R. B. Chung, C. Han, C. Pan, N. Pfaff, J. S. Speck, S. P. DenBaars, and S. Nakamura. *Appl. Phys. Lett.* **101** 131113 (2012).

[29] M. L. Nakarmi, K. H. Kim, K. Zhu, J. Y. Lin, and H. X. Jiang. *Appl. Phys. Lett.* **85** 3769 (2004).

[30] B. Jahnen, M. Albrecht, W. Dorsch, S. Christiansen, H. P. Strunck, D. Hanser, and R. F. Davis. *MRS Internet J. Nitride Semicond. Res.* **3** 39 (1998).

[31] T. Mattila and R. M. Nieminen. *Phys. Rev. B* **55** 9571 (1997).

[32] C. Stampfl and C. G. Van de Walle. *Appl. Phys. Lett.* **72** 459 (1998).

[33] C. H. Park and D. J. Chadi. *Phys. Rev. B* **55** 12995 (1997).

[34] R. Collazo, S. Mita J. Xie, A. Rice J. Tweedie, R. Dalmau, and Z. Sitar. *Phys. Status Solidi (c)* **8** 2031 (2010).

[35] L. Gordon, J. B. Varley, J. L. Lyons, A. Janotti, and C. G. Van de Walle. *In preparation.*

[36] L. Gordon, J. L. Lyons, A. Janotti, and C. G. Van de Walle. *In preparation.*

[37] O. Madelung. *Semiconductors–Basic Data: 2nd edition*. Springer and Berlin (1996).

[38] M. D. McCluskey, N. M. Johnson, C. G. Van de Walle, D. P. Bour, M. Kneissl, and W. Walukiewicz. *Phys. Rev. Lett.* **80** 4008 (1997).

[39] P. Mooney. *Deep Centers in Semiconductors*. Gordon, Breach, and Yverdon (1992).

[40] D. J. Chadi and K. J. Chang. *Phys. Rev. Lett.* **61** 873 (1989).

[41] M. L. Nakarmi, N. Nepal, J. Y. Lin, and H. X. Jiang. *Appl. Phys. Lett.* **86** 261902 (2005).

[42] M. Pophristic, S. P. Guo, and B. Peres. *Appl. Phys. Lett.* **82** 4289 (2003).

[43] P. Cantu, S. Keller, U. K. Mishra, and S. P. DenBaars. *Appl. Phys. Lett.* **82** 3683 (2003).

[44] C. H. Park and D. J. Chadi. *Phys. Rev. B* **55** 12995 (1997).

[45] P. Boguslawski and J. Bernholc. *Phys. Rev. B* **56** 9496 (1997).

[46] N. T. Son, M. Bickermann, and E. Janzen. *Appl. Phys. Lett.* **98** 092104 (2011).

[47] R. Zeisel, M. W. Bayerl, S. T. B. Goennenwein, R. Dimitrov, O. Ambacher, M. S. Brandt, and M. Stutzmann. *Phys. Rev. B* **61** R16283 (2000).

[48] X. T. Trinh, N. Nilsson, I. G. Ivanov, E. Janzen, A. Kakanakova-Georgieva, and N. T. Son. *Appl. Phys. Lett.* **103** 042101 (2013).

[49] C. Skierbiszewski, T. Suski, M. Leszczynski, M. Shin, M. Skowronski, M. D. Bremser, and R. F. Davis. *Appl. Phys. Lett.* **74** 3833 (1999).

[50] I. Vurgaftman and J. R. Meyer. *J. Appl. Phys.* **94** 3675 (2003).

[51] L. Silvestri, K. Dunn, S. Prawer, and F. Ladouceur. *Appl. Phys. Lett.* **99** 122109 (2011).

[52] R. Collazo, S. Mita J. Xie, A. Rice J. Tweedie, R. Dalmau, and Z. Sitar. *Phys. Status Solidi (c)* **8** 2031 (2011).

[53] S. Fritze, A. Dadgar, H. Witte, M. Bugler, A. Rohrbeck J. Blasing. A. Hoffmann, and A. Krost. *Appl. Phys. Lett.* **100** 122104 (2012).

[54] A. Harrison. *Elementary Electronic Structure*. World Scientific (1999).

[55] J. F. Carlin *et al*. *Phys. Status Solidi. B.* **242** 2326 (2005).

[56] R. Butte *et al*. *J. Phys. D.* **40** 6328 (2007).

[57] S. Y. Karpov, N. I. Podolskaya, I. A. Zhmakin, and A. I. Zhmakin. *Phys. Rev. B* **70** 235203 (2004).

[58] L. Gordon, H. Abu Farsakh, A. Janotti, and C. G Van de Walle. *Submitted*.

[59] L. Gordon J. R. Weber J. B. Varley, A. Janotti, D. D. Awschalom, and C G. Van de Walle. *MRS Bulletin* **38** 10 (2013).

[60] L. Gordon, A. Janotti, and C. G. Van de Walle.. *In preparation.*

[61] L. Gordon, A. Falk, A. Alkauskas, A. Janotti, and C. G Van de Walle. *In preparation.*

[62] W. A. Philips. *Rep. Prog. Phys.* **50** 1657 (1987).

[63] A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger. *Rev. Mod. Phys.* **59** 1 (1987).

[64] M. Steffen. *Physics* **4** 103 (1987).

[65] Y. Shalibo, Y. Rofe, D. Shwa, F. Zeides, M. Neeley, J. M. Martinis, and N. Katz. *Phys. Rev. Lett.* **105** 177001 (2010).

[66] E. Collin et al. *Phys. Rev. Lett.* **93** 157005 (2004).

[67] O. Astafiev et al. *Phys. Rev. Lett.* **93** 267007 (2004).

[68] J. M. Martinis et al. *Phys. Rev. Lett.* **95** 210505 (2005).

[69] R. O. Pohl. *Rev. Mod. Phys.* **74** 991 (2002).

[70] J. Gao et al. *Appl. Phys. Lett.* **92** 152505 (2008).

[71] S. Oh et al. *Phys. Rev. B* **74** 100502 (2006).

[72] A. M. Holder, K. D. Osborn, C. J. Lobb, and C. B. Musgrave. *Phys. Rev. Lett.* **111** 065901 (2013).

[73] C. G. Van de Walle and J. Neugebauer. *Nature* **423** 626 (2003).

[74] A. Novak. *Structure and Bonding* **18** 177 (1974).

[75] W. M. Hlaing Oo, S. Tabatabaei, M. D. McCluskey, J. B. Varley, A. Janotti, and C. G. Van de Walle. *Phys. Rev. B* **82** 193201 (2010).

[76] E. J. Spahr *et al*. *Phys. Rev. Lett.* **102** 075506 (2009).

[77] Z. Wang, T. Liu C. Li, and T. Sham. *Journal of Physics: Conference Series* **430** 012065 (2013).

[78] C. J. Ball. *Powder Diffraction* **21** 1 (2006).

[79] P. Vashishta, R. K. Kalia, A. Nakano, and J. Pedro Rino. *J. Appl. Phys.* **103** 083504 (2008).

[80] S. Davis and G. Gutierrez. *J. Phys. Condens. Matter* **23** 495401 (2011.

[81] T. C. DuBois, M. C. Per, S. P. Russo, and J. H. Cole. *Phys. Rev. Lett.* **110** 077002 (2013).

[82] R. Barends *et al. Phys. Rev. Lett.* **111** 080502 (2013).

[83] C. P. Flynn and A. M. Stoneham. *Phys. Rev. B* **1** 3966 (1970).

[84] P. G. Sundell and G. Wahnstrom. *Phys. Rev. Lett.* **92** 155901 (2004).

[85] B. G. Johnson, P. M. W. Gill, and J. A. Pople. *J. Chem. Phys.* **98** 5612 (1993).

[86] M. Choi, J. L. Lyons, A. Janotti, and C. G. Van de Walle. *Phys. Status Solidi b.* **250** 787 (2013).

[87] P. G. Sundell, M. E. Bjorketun, and G. Wahnstrom. *Phys. Rev. B* **76** 096301 (2007).

[88] J. H. Cole, C. Muller, P. Bushev, G. J. Grabovskij, J. Lisenfeld, A. Lukashenko, A. V. Ustinov, and A. Ahnirman. *Appl. Phys. Lett.* **97** 252201 (2010).

[89] M. S. Khalil *et al*. *Appl. Phys. Lett.* **103** 162601 (2013).

[90] G. Pilania *et al*. *Scientific Reports* **4** 4485 (2014).

[91] P. Neumann, N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko, and J. Wrachtrup. *Science* **320** 1326 (2008).

[92] J. R. Weber, W. F. Koehl J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle, and D. D. Awschalom. *Proc. Nat. Acad. Sci.* **107** 8513 (2010).

[93] N. B. Manson, J. P. Harrison, and M. J. Sellars. *Phys. Rev. B* **74** 104303 (2006).

[94] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markhan, R. Kolesov, N. Mizuochi J. Isoya J. Achard J. Beck J. Tissler, V. Jacques, P. R. Hemmer, F. Jelezko, and J. Wrachthrup. *Nat. Mater.* **8** 383 (2009).

[95] G. Davies and M. F. Hamer. *Proc. R. Soc. Lond. Ser. A.* **348** 285 (1976).

[96] A. Gali. *Phys. Status Solidi (b)* **248** 1337 (2011).

[97] W. Koehl, B. B. Buckley, F. J. Heremans, G. Calusine, and D. D. Awschalom. *Nature* **479** 84 (2011).

[98] A. L. Falk, B. B. Buckley, G. Calusine, W. F. Koehl, V. V. Dobrovitski, A. Politi, C. A. Zorman, P. X.-L. Feng, and D. D. Awschalom. *Nature Comm.* **4** 1819 (2013).

[99] A. L. Falk, P. V. Klimov, B. B. Buckley, V Ivady, I. A. Abrikosov, G. Calusine, Koehl. W. F., A. Gali, and D. D. Awschalom. *Phys. Rev. Lett.* **112** 187601 (2014).

[100] D. J. Christle, A. L. Falk, P. Andrich, P. V. Klimov, J. Hassan, N. T. Son, E. Janzn, T. Ohshima, and D. D. Awschalom. *arXiv:1406.7325* (2014).

[101] V.A. Soltamov, A.A. Soltamova, P.G. Baranov, and I.I. Proskuryakov. *Phys. Rev. Lett.*, **108** 226402 (2012).

[102] A. Powell J. Jenny, S. Muller, H. McD. Hobgood, V. Tsvetkov, R. Lenoard, and C. Carter. *Int. J. High Speed Electron. Syst.* **16** 751 (2006).

[103] C. M. Zetterling (ed.) *Process Technology for Silicon Carbide Devices*. Institute of Engineering and Technology (2002).

[104] N. T. Son, P. Carlsson J. ul Hassan, E. Janzén, T. Umeda J. Isoya, A. Gali, M. Bockstedte, N. Morishita, T. Ohshima, and H. Itoh. *Phys. Rev. Lett.* **96** 055501 (2006).

[105] A. J. Steckl J. Devrajan, S. Tlali, H. E. Jackson, C. Tran, S. N. Gorin, and L. M. Ivanova. *Appl. Phys. Lett.* **69** 3824 (1996).

[106] C. A. Zorman, A. J. Fleischman, A. S. Dewa, M. Mehregany, C. Jacob, S. Nishino, and P. Pirouz. *J. Appl. Phys.* **78** 5136 (1995).

[107] Y. Goldberg, M. E. Levinshtein, and S. L. Rumyantsev. *Properties of Advanced Semiconductor Materials*. Wiley (2001).

[108] L. Torpo, T. E. M. Staab, and R. M. Nieminen. *Phys. Rev. B* **65** 8 085202 (2002).

[109] N. T. Son, E. Sörman, W. M. Chen, C. Hallin, O. Kordina, B. Moneman, E. Janzen, and J. L. Lindström. *Phys. Rev. B* **55** 2863 (1997).

[110] A. Taylor and R. M. Jones. *Silicon Carbide - A High Temperature Semiconductor*. Pergamon Press (1960).

[111] A. Alkauskas, J. Lyons, D. Steiauf, , and C. G. Van de Walle. *Phys. Rev. Lett.* **109** 267401 (2012).

[112] J. Schneider, H. D. Muller, K. Maier, W. Wilkening, F. Fuchs, A. Dornen, S. Leibenzeder, and R. Stein. *Appl. Phys. Lett.* **56** 1184 (1990).

[113] R. A. Babunts, V. A. Vetrov, I. V. Ivin, E. N. Mokhov, N. G. Romanov, V. A. Khramtsov, and P. G. Baranov. *Physics of the Solid State* **56** 1184 (1999).

[114] J. R. Weber, W. F. Koehl J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle, and D. D. Awschalom. *J. Appl. Phys.* **109** 102417 (2011).

[115] M. Fehr, A. Schnegg, B. Rech, K. Lips, O. Astakhov, F. Finger, G. Pfanner, C. Freysoldt, J. Neugebauer, R. Bittl, and C. Teutloff. *Phys. Rev. B* **84** 245203 (2011).

[116] C. G. Van de Walle, D. B. Laks, G. F. Neumark, and S. T Pantelides. *Phys. Rev. B.* **47** 9425 (1993).

[117] T. Chanier, C.E. Pryor, and M.E. Flatte. *Europhys. Lett.*, **99** 67006 (2012).

[118] X. Wang, M. Zhao, Z. Wang, X. He, Y. Xi, and S. Yan. *Appl. Phys. Lett.*, **100** 192401 (2012).

[119] D.H. Tanimoro, W.M. Ziniker, and J.O. Kemp. *Phys. Rev. Lett.*, **14** 645 (1965).

[120] B. Henderson. *J. Phys. C*, **9** 2185 (1976).

[121] I. Aharonovich, S. Castelletto, B.C. Johnson, J.C. McCallum, D.A. Simpson, A.D. Greentree, and S. Prawer. *Phys. Rev. B*, **81** 121201 (2010).

[122] S. Bertaina, L. Chen, N. Groll, J. Van Tol, N.S. Dalal, and I. Chiorescu. *Phys. Rev. Lett.*, **102** 050501 (2009).

[123] A. Janotti, L. Bjaalie, L. Gordon, and C. G. Van de Walle. *Phys. Rev. B* **86** 241108(R) (2012).

[124] L. Gordon, A. Sarwe, D. Steiauf, A. Janotti, and C. G Van de Walle. *In preparation.*

[125] B. Himmetoglu, A. Janotti, K. Krishnaswamy, L. Bjaalie, L. Gordon, and C. G. Van de Walle. *In preparation.*

[126] L. Gordon, M. Miao, S. Chowdhury, M. Higashiwaki, U. K. Mishra, and C. G. Van de Walle. *Journal of Physics D: Applied Physics* **43** 505501 (2010).

[127] J. L. Lyons, L. Gordon, Anderson Janotti, and C. G. Van de Walle. *In preparation.*

[128] A. Ohtomo and H. Hwang.. *Nature* **427** 423 (2004).

[129] J. Son, P. Moetakef, B. Jalan, O. Bierwagen, N. J. Wright, R. Engel-Herbert, and S. Stemmer. *Nature Mater.* **9** 482 (2010).

[130] Y. Hotta, T. Susaki, and H. Y. Hwang. *Phys. Rev. Lett.* **99** 236805 (2007).

[131] H. W. Jang, D. A. Felker, C. W. Bark, Y. Wang, M. K. Niranjan, C. T. Nelson, Y. Zhang, D. Su, C. M. Folkman, S. H. Baek, S. Lee, K. Janicka, Y. Zhu, X. Q. Pan, D. D. Fong, E. Y. Tsymbal, M. S. Rzchowski, and C. B. Eom. *Science* **331** 886 (2011).

[132] L. Bjaalie, B. Himmetoglu, L. Weston, A. Janotti, and C. G. Van de Walle. *New Journal of Physics* **16** 025005 (2014).

[133] S. Thiel, G. Hammerl, A. Schmehl, C. W. Schneider, and J. Mannhart. *Science* **313** 1942 (2006).

[134] C. Cen, S. Thiel, J. Mannhart, and J. Levy. *Science* **323** 1026 (2009).

[135] B. Forg, C. Richter, and J. Mannhart. *Appl. Phys. Lett.* **100** 053506 (2012).

[136] F. Chudnovskiy, S. Luryi, and B. Spivak. *Future Trends in Microelectronics: The Nano Millennium*. Wiley (2002).

[137] J. Mannhart, D. H. A. Blank, H. Y. Hwang, A. J. Millis, and J. M. Triscone. *MRS Bull.* **33** 1027 (2008).

[138] A. Huijben, G. Brinkman, G. Koster, G. Rijnders, H. Hilgenkamp, and D. H. A. Blank. *Adv. Mater.* **21** 1665 (2009).

[139] A. Brinkman, M. Huijben, M. Van Zalk, M. Huijben, U. Zeitler, J. C. Maan, W. G. Van der Wiel, G. Rijnders, D. H. A. Blank, and H. Hilgencamp. *Nature Mater.* **6** 493 (2007).

[140] N. Reyren, S. Thiel, A. D. Caviglia, L. Kourkoutis, G. Hammerl, C. Ricther, C. W. Schneider, T. Kopp, A. S. Ruetschi, D. Jaccard, M. Gabay, D. A. Muller, J. M. Triscone, and J. Mannhart. *Science* **317** 1196 (2007).

[141] L. Li, C. Richter, J. Mannhart, and R. C. Ashoori. *Nature Phys.* **7** 762 (2011).

[142] H. Y. Hwang, Y. Iwasa, M. Kawasaki, B. Keimer, N. Nagaosa, and Y. Tokura. *Nature Mater.* **11** 103 (2012).

[143] M. S. Park, S. H. Rhim, and A. J. Freeman. *Phys. Rev. B* **74** 205416 (2006).

[144] W. Siemons, G. Koster, H. Yamamoto, W. A. Harrison, G. Lucovsky, T. H. Geballe, D. H. A. Blank, and M. R. Beasley. *Phys. Rev. Lett.* **98** 196802 (2007).

[145] J. Lee and A. A. Demkov. *Phys. Rev. B* **78** 193104 (2008).

[146] R. Pentcheva and W. E. Pickett. *J. Phys.: Condens. Matter* **22** 043001 (2010).

[147] N. C Bristowe, P. B. Littlewood, and E. Artacho. *Phys. Rev. B* **83** 205405 (2011).

[148] Y. Li, S. N. Phattalung, S. Limpijumnong, J. Kim, and J. Yu. *Phys. Rev. B* **84** 245307 (2011).

[149] N. Pavlenko, T. Kopp, E. Y. Tsymbal, G. A. Sawatzky, and J. Mannhart. *J. Phys. Rev. B* **85** 020407 (2012).

[150] M. Basletic, Carrétéro Mauric J.-L., C., G. Herranz, O. Copie, M. Bibes, E. Jacquet, K. Bouzehouane, S. Fusil, and A. Barthélémy. *Nature Mater.* **7** 621 (2008).

[151] A. Kalabukhov, R. Gunnarsson, J. Börjesson, E. Olsson, T. Claeson, and D. Winkler. *Phys. Rev. B* **75** 121404(R) (2007).

[152] S. Thiel, G. Hammerl, A. Schmehl, C. W. Schneider, and J. Mannhart. *Science* **313** 1942 (2006).

[153] G. Herranz, M. Basletic, M. Bibes, C. Carrétéro, E. Tafra, E. Jacquet, K. Bouzehouane, C. Deranlot, A. Hamzic, J. M. Broto, A. Barthélémy, and A. Fert. *Phys. Rev. Lett.* **98** 216803 (2007).

[154] M. Huijben, G. Rijnders, D. H. A. Blank, S. Bals, S. Van Aert, Verbeeck J., G. Van Tendeloo, A. Brinkman, and H. Hilgenkamp. *Nature Mater.* **5** 556 (2006).

[155] N. Nakagawa, H. Y. Hwang, and D. A. Muller. *Nature Mater.* **5** 204 (2006).

[156] P. R. Willmott, S. A. Pauli, R. Herger, C. M. Schlepütz, D. Martoccia, B. D. Patterson, B. Delley, R. Clarke, D. Kumah, C. Cionca, and Y. Yacoby. *Phys. Rev. Lett.* **99** 155502 (2007).

[157] L. Qiao, T. Droubay, V. Shutthanandan, Z. Zhu, P. Sushko, and S. J. Chambers. *Phys. Condens. Matter* **22** 312201 (2010).

[158] P. Moetakef, T. A. Cain, D. G. Ouellette, J. Y. Zhang, D. O. Klenov, A. Janotti, C. G. Van de Walle, S. Rajan, S. J. Allen, and S. Stemmer. *Appl. Phys. Lett.* **99** 232116 (2011).

[159] W. Son, E. Cho, Lee J., and S. Han. *J. Phys. Cond. Matt.* **22** 315501 (2010).

[160] R. Arras, V. G. Ruiz, W. E. Pickett, and R. Pentcheva. *Phys. Rev. B* **85** 125404 (2012).

[161] R. A. van der Berg, P. W. M. Blom, J. F. M. Cillessen, and R. M. Wolf. *Appl. Phys. Lett.* **66** 6 (1995).

[162] X. Lu, Z. Liu, Y. Wang, Y. Yang, X. Wang., H. Zhou., and B. Nguyen. *J. Appl. Phys.* **94** 1229 (2003).

[163] S. Rajan, P. Waltereit, C. Poblenz, S. J. Heikman, D. S. Green, J. S. Speck, and U. K. Mishra. *IEEE Electron Device Lett.* **25** 247 (2004).

[164] E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau. *Appl. Phys. Lett.* **71** 2794 (1997).

[165] O. Ambacher J. Smart J. R. Shealy, N. G. Weimann, K. Chu, M. Murphy, W. J. Schaff, L. F. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger, and J. Hilsenbeck *J. Appl. Phys.* **85** 3222 (1999).

[166] I. P. Smorchkova, C. R. Elsass J. P. Ibbetson, R. Vetury, B. Heying, P. Fini, E. Haus, S. P. DenBaars, J. S. Speck, and U. K. Mishra. *J. Appl. Phys.* **86** 4520 (1999).

[167] J. P. Ibbetson, P. T. Fini, K. D. Ness, S. P. DenBaars, J. S. Speck, and U. K. Mishra. *Appl. Phys. Lett.* **77** 250 (2000).

[168] I. P. Smorchkova, L. Chen, T. Mates, L. Shen, S. Heikman, B. Moran, S. Keller, S. P. DenBaars, J. S. Speck, and U. K. Mishra. *J. Appl. Phys.* **90** 5196 (2001).

[169] S. Heikman, S. Keller, Y. Wu, J. S. Speck, S. P. DenBaars, and U. K. Mishra. *J. Appl. Phys.* **93** 10114 (2003).

[170] G. Koley and M. G. Spencer. *Appl. Phys. Lett.* **86** 042107 (2005).

[171] S. Chowdhury, M. Higashiwaki, M. S. Miao, B. L. Swenson, C. G. Van de Walle, and U. K. Mishra. *36th International Symposium on Compound Semiconductors* (2009).

[172] G. Koley, H. Y. Cha, J. Hwang, W. J. Schaff, L. F. Eastman, and M. G. Spencer. *J. Appl. Phys.* **96** 4253 (2004).

[173] M. Higashiwaki, S. Chowdhury, M. S. Miao, B. L. Swenson, C. G. Van de Walle, and U. K. Mishra. *J. Appl. Phys.* **108** 6 (2010).

[174] J. A. Majewski, S. Hackenbuchner, G. Zandler, and P. Vogl. *Comput. Mater. Sci.* **30** 81 (2004).

[175] J. M. Tirado, J. L. Sanchez-Rojas, and J. I. Izpura. *Semicond. Sci. and Technol.* **21** 1150 (2006).

[176] M. Miczek, B. Adamowicz, C. Mizue, and T. Hashizume. *Jpn. J. Appl. Phys.* **48** 04C092 (2009).

[177] C. Morkoc, J. Leach, C. Wood, and D. Jena. *Polarization effects in semiconductors*. Springer (2007).

[178] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan. *J. Appl. Phys.* **89** 5815 (2003).

[179] I. Vurgaftman and J. R. Meyer. *J. Appl. Phys.* **94** 3675 (2003).

[180] M. S. Miao, Anderson Janotti, and C. G. Van de Walle. *Phys. Rev. B* **80** 155319 (2009).

[181] C. G. Van de Walle and D. Segev. *J. Appl. Phys.* **101** 081704 (2007).

[182] M. Miao, P. G. Moses J. R. Weber, A. Janotti, and C. G. Van de Walle. *Europhys. Lett.* (2010).

[183] M. Miao, J. R. Weber, and C. G. Van de Walle. *J. Appl. Phys.* **107** 12 (2010).

[184] J. Simon, V. Protasenko, C. Lian, and D. Jena. *Science* **327** 6064 (2010).

[185] S. Rajan, A. Chini, M. H. Wong, J. S. Speck, and U. K. Mishra. *J. Appl. Phys.* **104** 4 (2007).

[186] H. Wong, U. Singisetti, J. Lu, J. S. Speck, and U. K. Mishra. *IEEE Trans. on Elec. Devices* **59** 11 (2012).

[187] W. Kaun, P. G. Burke, M. H. Wong, E. C. H. Kyle, U. K. Mishra, and J. S. Speck. *Appl. Phys. Lett.* **101** 26 (2012).